INVESTIGATION OF THE CHARACTERISTICS OF THE SURFACE-POTENTIAL CONTROLLED TRANSISTOR

PHILLIP B. FRIEDRICHS

INVESTIGATION OF THE CHARACTERISTICS OF THE SURFACE-POTENTIAL CONTROLLED TRANSISTOR

by

Phillip B. Friedrichs

Captain, United States Marine Corps

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE
IN
ENGINEERING ELECTRONICS

United States Naval Postgraduate School
Monterey, California

1 9 6 2

INVESTIGATION OF THE CHARACTERISTICS OF THE SURFACE-POTENTIAL CONTROLLED TRANSISTOR

by

Phillip B. Friedrichs

This work is accepted as fulfilling

the thesis requirements for the degree of

MASTER OF SCIENCE

IN

ENGINEERING ELECTRONICS

from the

United States Naval Postgraduate School

ABSTRACT

A surface potential controlled transistor was recently developed, the characteristics of which depend upon two semiconductor phenomena usually thought of as deleterious to transistor operation. These are carrier recombination and channel effects. The device is characterized by a high impedance grid which allows efficient operation from a low impedance source. It offers promise of replacing vacuum tubes in many applications not previously possible.

The physics of operation are explained and the description of the device in terms of hybrid parameters is given. The variation of these parameters with bias conditions is investigated and explained in terms of the underlying physics. The characterization of the device as a transducer is derived in terms of gain and impedance levels.

The writer wishes to express his appreciation to Mr. James Martin and Mr. Ben Anixter of Fairchild Semiconductor Products; to Mr. Fred Morrison and Mr. Randolph Moore of Motorola, Inc., Military Electronics Division for their assistance in obtaining the devices for experimentation; and to Dr. W. M. Bauer of the U.S. Naval Postgraduate School for his assistance and encouragement in this investigation.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

1. Introduction.

In 1961, Dr. C.T. Sah described a new semiconductor device which he called a "surface-potential controlled transistor". /1/ In addition to the usual emitter, base and collector terminals this device features a fourth terminal known as a grid. It is the effect on collector current of a potential applied to the grid and the high impedance of this terminal which sets this device apart from the remainder of the transistor family. Because of the fourth terminal and the similarities to vacuum tube operation, this device is appropriately described by the term "tetrode" and such it shall be called during the remainder of this discussion.

The tetrode depends upon two semiconductor phenomena for its operation. These phenomena are channel and recombination of holes and electrons. Until the advent of the tetrode these two effects were deleterious to transistor action and strong efforts were made to curb them. Since they were regarded as second order effects, most engineers are unfamiliar with the mechanisms involved. They will be described in Sections 3 and 4 of this paper.

The tetrode may be described by a set of three hybrid equations defined by the current-voltage relationships at each of the three independent terminals. These characteristics and their variations with bias conditions will be presented in Section 7 along with the associated linear equivalent circuit. An explanation of the parameters in terms of the physics of the tetrode will also be given. Finally, the expressions for voltage gain, input impedance and output impedance will be derived and discussed together with some cautions regarding use of this device in circuits.

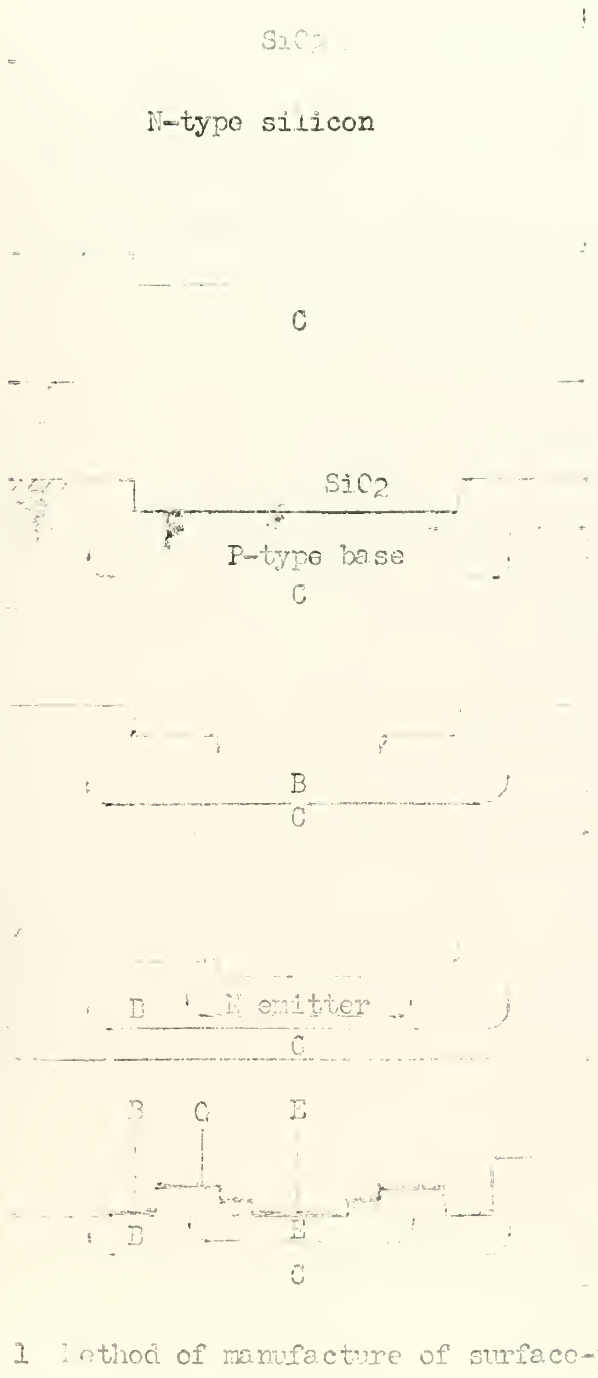
2. Description of Manufacture.

A knowledge of the construction of the tetrode will be an aid in understanding the action of the transistor in later discussions. Construction follows the sequence for the surface passivated planar. However, the last phase involves depositing a metal contact on the protective oxide layer of the emitter-base junction. Figure 1 illustrates the steps in manufacture.

An n-type silicon substrate forms the collector. This is etched to clean the surface and then exposed to a steam and oxygen atmosphere which causes a stable silicon dioxide to grow into the surface. It is this protective oxide which passivates the surface and provides the extremely low leakage currents in planar transistors (of the order of tenths of nanoamps). This is shown in figure 1(a).

A window is etched in the oxide layer exposing the n-type silicon for the base diffusion (figure 1(b)). The base is diffused into the collector from a boric acid vapor atmosphere. The metallic boron diffuses into the exposed silicon and forms a p-type layer. It is important to note in figure 1(c) that the junction is formed underneath the oxide layer covering the collector. This means that the junction surface has not been exposed to a contaminating atmosphere as is the case in conventional transistor manufacture. During the base diffusion, another oxide layer grows into the previously exposed silicon. In the case where boron is used as the base diffusant, this may take the form of a borosilicate glass.

As before, a window is etched in the oxide over the base to expose the p-type silicon to the emitter diffusant (figure 1(d)). Phosphorous is used as the donor impurity for the emitter n-type layer. It is diffused into the base from a phosphorous pentoxide atmosphere. The



(a) Oxide grows into entire substrate

(b) Window etched for base diffusion

(c) Oxide grows deeper during boron diffusion of base

(d) Window etched for emitter diffusion

(e) Oxide grows deeper during phosphorous diffusion of emitter

(f) Windows for ohmic contacts etched. Emitter, base and grid oxide are metallized

Fig. 1 Method of manufacture of surface-potential controlled transistor.

presence of the oxygen causes a third oxide to grow into the exposed silicon and again, the junction intersects the surface under the protective oxide layer which passivates the surface. (figure 1(e))

Provision for ohmic contacts is made by etching windows over the emitter and base regions and metallizing with aluminum. The etching of the ohmic contact windows leaves an oxide layer protecting the emitter-base junction surface. It is this oxide which provides the possibility of a fourth terminal. This oxide is metallized and a lead is bonded to it, thus forming the grid. The effect of a potential applied to this grid is the subject of the remainder of this paper.

3. The Phenomenon of Channel in Semiconductors.

An understanding of the way in which collector current is controlled by grid voltage in the tetrode requires an understanding of an effect known as channel conductivity in semiconductors. This effect has been observed for many years in transistors and has been used to explain excessive leakage current in p-n junctions /2/ as well as the drift in potential of an open-circuited emitter when the collector-base junction is reverse biased./3/ This latter effect is known as "floating potential."

Channel conductivity may be described as a region near a p-n junction into which either electrons or holes from the other side of the junction may flow without encountering a potential barrier./1/ This is often mistaken for an inversion layer which is a layer of n-type semiconductor near the surface of a material which has been doped to be p-type or vice versa. Several investigators have concerned themselves with methods of creating and controlling channels artificially in the laboratory. J.T. Law attributed this anomalous conductivity to an ionic current in a layer of water adsorbed on the surface when a reverse bias was applied to a p-n junction./4/ At approximately the same time, H. Christensen isolated the channel conductivity from Law's ionic current theory by freezing the adsorbed water layer at dry ice temperature. He was thus able to show that the channel was, at least in part, due to charge carriers within the bulk./5/ He attributed the channel to an electric field at the surface. This electric field disturbed the equilibrium charge distribution near the surface thus causing the channel. W.L. Brown /3/, following J. Bardeen's theory of surface states /6/ developed a mathematical treatment for describing the electric field in the interior and at the surface of a semiconductor with such an adsorbed layer of surface charge. His model was extended by A.L. McWhorter and R.H. Kingston who assumed a

channel and calculated the associated currents and voltages in and along the channel for a germanium model./2/ As Sah points out /1/, this model is valid only for germanium samples since recombination current was neglected. Such recombination is of the utmost importance in describing the tetrode and will be discussed in detail later. In 1957, M. Cutler and H.M. Bath modified the work of McWhorter and Kingston to explain the non-saturating reverse current and low forward current characteristics in silicon diodes./7/ R.H. Kingston and S.F. Neustadter followed a development parallel to Brown's in calculating the space charge and free carrier concentration as well as the electric field at the surface of a semiconductor./8/ These works provide a useful departure point for the treatment of the channel effects in the Sah tetrode. Unfortunately, all of these treatments are directed toward the individual who is well versed in the concepts of the physics of the solid state and not the average electronics engineer. The development which follows is a model for the creation of a channel which uses concepts familiar to the engineer. The reader should be cautioned against pressing the model beyond its applicability. It is intended as a heuristic model. A more comprehensive model within the precepts of solid state physics will be presented later. The reader who is interested in following the later development or who feels need of a review of semiconductor physics is referred to the now classic book on the subject by W. Shockley; particularly chapters 1, 5, 9, 10 and 12. A foundation in quantum mechanics oriented toward semiconductors is set out in Part III of the same volume./9/

As stated earlier, there is a difference between the formation of a channel and an inversion layer. A channel is associated only with the problem of whether a charge carrier sees a potential barrier at a junction, regardless of the actual carrier concentrations in the immediate

vicinity. An inversion layer is defined in terms of the relative carrier concentrations and may occur independent of whether the carriers see a potential barrier. Inversion is present whenever the concentration of the normally minority carriers exceeds the normally majority carrier concentration. This may occur before, at the same time, or after the onset of a channel. A channel occurs whenever the potential barrier for majority carriers in a p-n junction is reduced to zero. For example, if electrons in an n-type material see no potential barrier to restrain them from diffusing to the p side of a p-n junction, the condition of channel is said to exist. To simplify matters, it will be assumed that the onset of channel and the formation of an inversion layer occur simultaneously. Under this assumption the words "channel" and "inversion layer" may be used interchangeably.

Consider the p-n junction in thermal equilibrium shown in figure 2(a). Under these conditions the junction is back biased by an amount V_d determined by the diffusion of majority carriers across the junction. The majority carrier space distribution is uniform in the y direction. The electrons on the n side now cannot surmount the repelling electric field in the transition region.

Now let a positive potential V_o be applied to the surface of the p-type material. The associated electric field repels the holes nearest the surface and uncovers some of the negative ions. Minority electrons are also attracted from the bulk toward the surface. The migration of the charge carriers and the uncovering of ions ends when the potential at the surface has been neutralized and the system is once more in equilibrium. One effect has been to reduce the concentration of free holes near the surface and increase the number of free electrons from their normal equilibrium values. Another effect has been to lower the potential

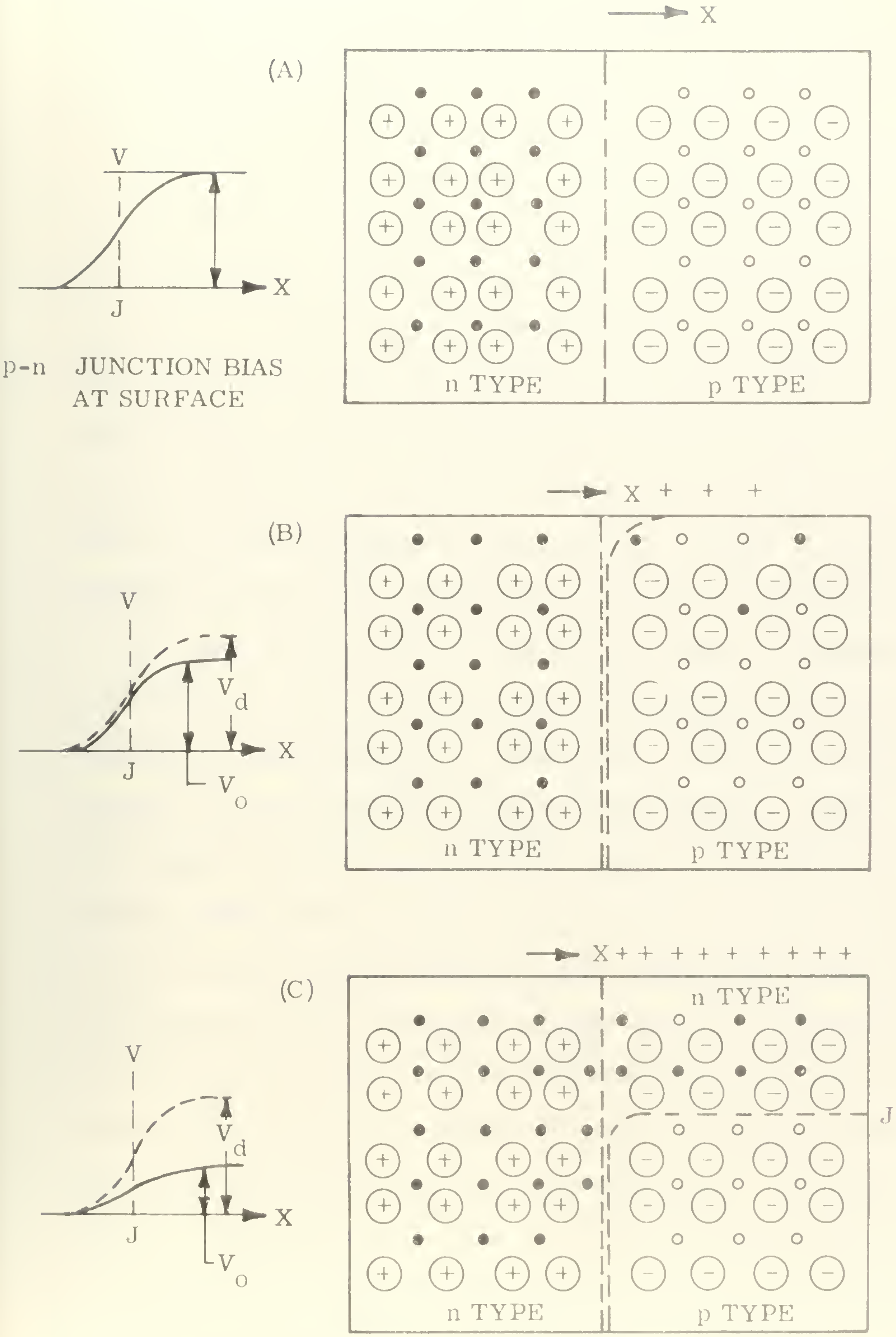


Fig. 2 Progressive steps in the formation of channel.

difference across the junction near the surface. Thus, a few of the neutralizing charges are electrons from the n side caused by this slight biasing. The new equilibrium condition is shown in figure 2(b). It is assumed that no current flows out through the semiconductor surface. The region near the surface is now less p-type than before.

Let the surface potential be increased. The number of electrons available from the p-type bulk is limited by the depth of penetration of the electric field into the bulk. However, increasing the surface potential tends to forward bias the junction near the surface and thus makes a large supply of electrons available in this area. Therefore, after an initial supply from the p-type material, nearly all of the electrons needed to neutralize the surface potential come from the n side and the change in surface potential is manifested in a change in bias across the junction near the surface necessary to provide these electrons. It should be clear that for some value of surface potential, the number of electrons will exceed the number of holes in the local region at the surface. This describes an n type material by definition. The further the potential increases, the more strongly n-type the layer becomes. Since the electric field caused by the surface potential decreases as a function of depth from the surface into the bulk, the n-type layer is less strongly n-type as we go from the surface into the bulk. Eventually, at some depth, the concentration of free holes again exceeds the free electron concentration and the material is again p-type. This indicates that the n-type layer does not extend very far into the bulk since the change in surface potential serves more to forward bias the junction than to repel holes from the surface. The fact that the depth of the surface n-type layer saturates very quickly will be demonstrated by the results of a later calculation. It should be noted that we have satisfied our definition of a channel;

an n-type layer in a region normally doped to be p-type.

It should be pointed out that the region where the bulk material makes the transition from the n-type surface layer, through the intrinsic range to the p-type region defines a p-n junction transition region. It should also be noted that this junction is reverse biased since the n region is at a positive potential with respect to the p region. To prepare the way for later arguments, consider the effect of a longitudinal current flowing in the channel toward the junction. The channel is quite resistive and a voltage drop occurs in the channel which reduces the reverse bias across the channel junction. Thus the channel becomes wider as we approach the normal junction from the nominally p side. As we shall see later, such a current is caused by recombination of electrons and holes in the channel. Figure 3 represents this situation for a constant current flowing in a channel of uniform resistivity. The potential drop is then a linear function of distance.

The case which is of interest to us is the tetrode. The source of the positive potential would, of course, be a source applied to the grid of the tetrode. Since the oxide layer over the emitter-base junction is an excellent insulator with a resistivity of approximately $5x10^9$ ohm-cm at 250°C the previous assumption of no current leaving the surface is valid. The description of channel formation used only a potential applied to the p-type material. The oxide layer which forms the grid extends over the n-type material as well. Such as applied potential on the n-side would nullify the biasing effects of the change in potential on the p-side. Thus it must be assumed that the nature of the interface between the silicon and the oxide is different for the two types of silicon and prevents the grid voltage from changing the potential at the active surface of the n-type silicon. Sah has attributed this effect to the

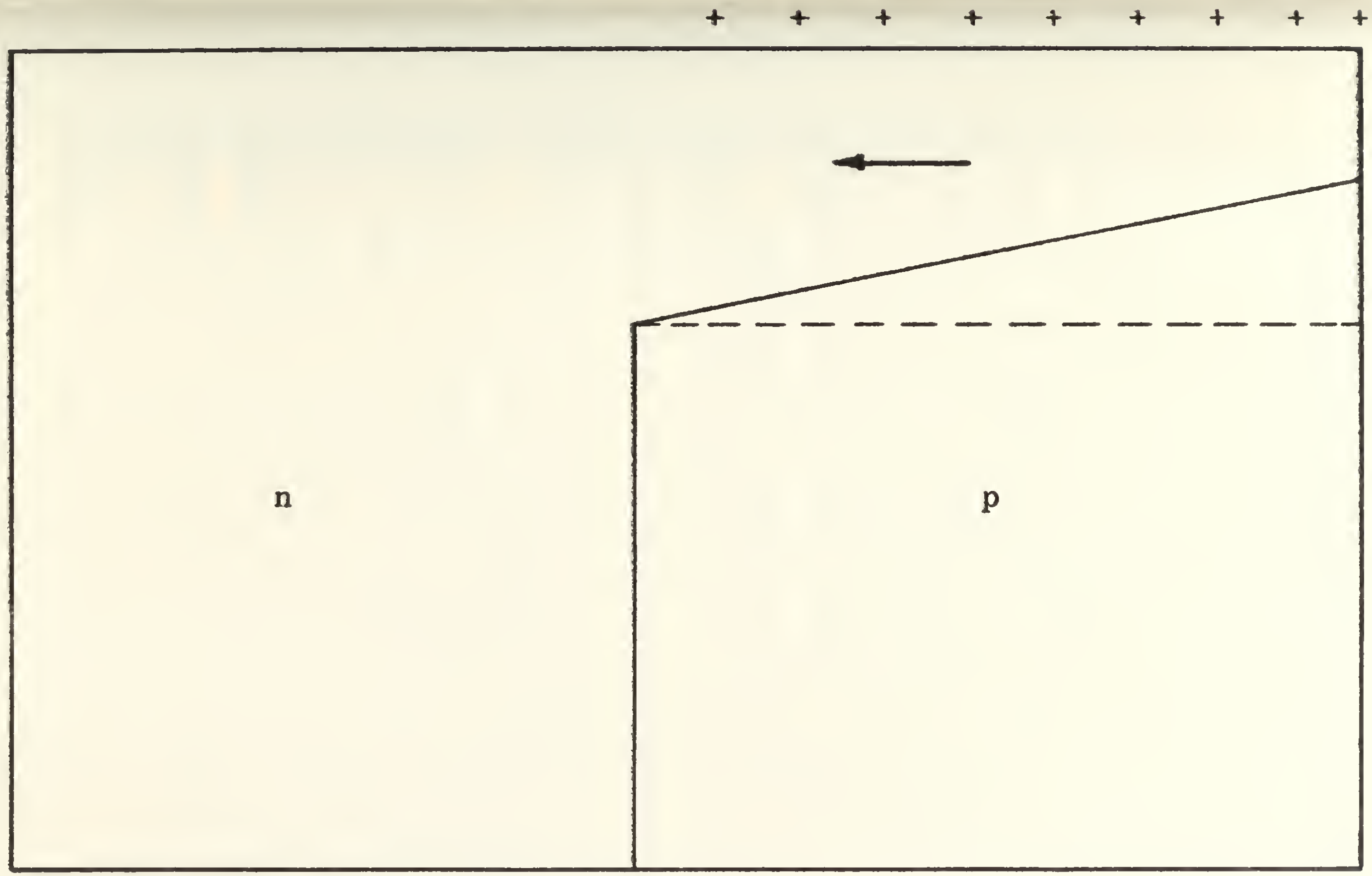


Fig. 3 Effect of channel current on channel depth.

probability that the phosphorous (n-type diffusant) is rejected by the oxide during the diffusion of the emitter and will therefore form a layer of highly concentrated phosphorous atoms at the interface. This forms a degenerate layer (metallic layer) which screens the surface of the silicon from the applied grid potential. Hence an applied grid voltage will not affect the distribution of charge carriers in the n-type material to any marked degree. On the other hand, the oxide shows an affinity for the base diffusant (boron) during diffusion so that the region at the surface, under the oxide is less heavily doped than the bulk material in the interior. As a result, the surface has a lower equilibrium hole concentration and the charge distribution will be more responsive to grid voltage changes. Figure 4 illustrates these points.

Formation of a channel on the p-side of an n-p junction has been described. This corresponds to a channel under the grid in the base of an n-p-n tetrode. In order to more fully describe the formation of the base channel, we must turn to an analysis of potential within the base.

The qualitative description of channel formation given above can be made more quantitatively rigorous by determining the electric field in the bulk of the base. More importantly, solution of the problem for the electrostatic potential within the volume will yield information as to the depth of penetration of the channel into the bulk. Later it will be possible to correlate potential with channel current by considering recombination current in the channel and its effect upon the depth of the channel. In order to formulate the problem, it will be necessary to describe channel formation in terms of the energy of charge carriers. The description will be somewhat parallel to that given by Brown /3/ and Kingston and Neustadter /8/.

Figure 5 is a three dimensional representation of the energy of an

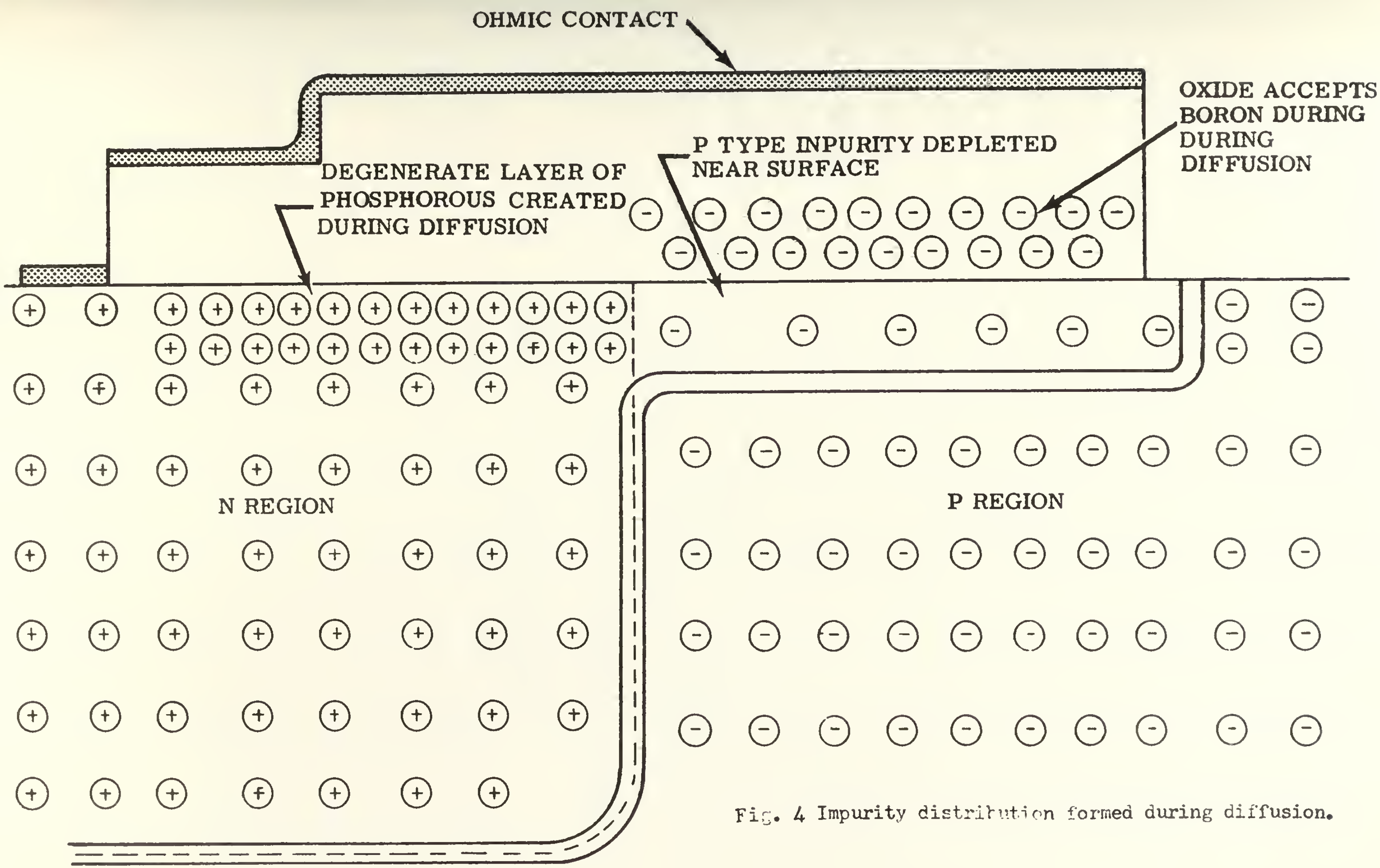


Fig. 4 Impurity distribution formed during diffusion.

electron as a function of depth into the bulk from the surface (denoted by y) and the distance across the surface from the n-type to the p-type material (x in this diagram). The silicon-oxide interface of the tetrode is represented by the plane $y = 0$ and the emitter and base regions are the n and p regions respectively. Figure 5(a) shows the energy levels of the junction at equilibrium. It is assumed that there is no residual deformation of the energy levels in the vicinity of the surface. ψ_p and ψ_n are the potentials at which the Fermi energy would lie for equal concentrations of electrons and holes. At the intersection of this surface and the Fermi surface, φ, $n = p$. For the equilibrium condition this occurs at the center of the junction. As before, it will be assumed that a potential applied to the grid (the plane $y = 0$) will not change the potential at the surface in the n region. V_0 is the reverse bias generated to counteract the diffusion of carriers across the junction. The subscript 0 refers to the thermal equilibrium value of the quantity concerned. The subscript n or p refers to the value of the quantity on the n or p side respectively.

The potential ψ_p or ψ_n is defined by the relationship

$$\psi = \frac{1}{2}\left[\psi_c + \psi_v + kT/q \ln (N_v/N_c)\right]$$

where q is the electronic charge, ψ_c and ψ_v are the potentials at the bottom of the conduction band and top of the valence band respectively, k is Boltzmann's constant, T is the ambient temperature in $^{\circ}$K and N_v and N_c are the effective densities of states in the conduction and valence bands respectively as defined by Shockley on page 303 of his book./9/ In a perfectly intrinsic material $N_v = N_c$ and $\psi = \frac{1}{2}(\psi_c + \psi_v)$ or one half of the energy gap.

In figure 5(b) a forward bias V_a has been applied to the junction in the conventional manner. This is manifested in a lowering of the

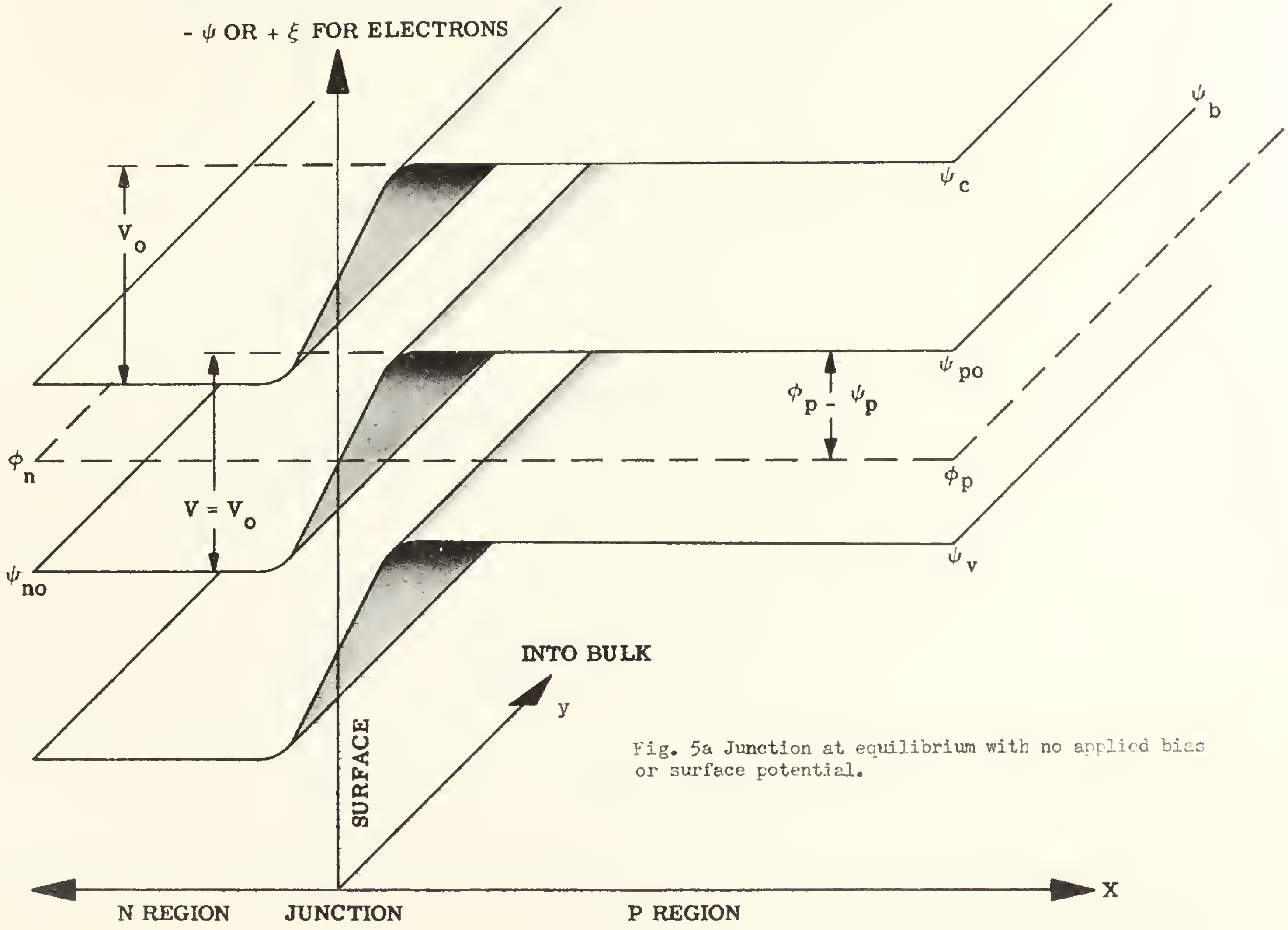


Fig. 5a Junction at equilibrium with no applied bias or surface potential.

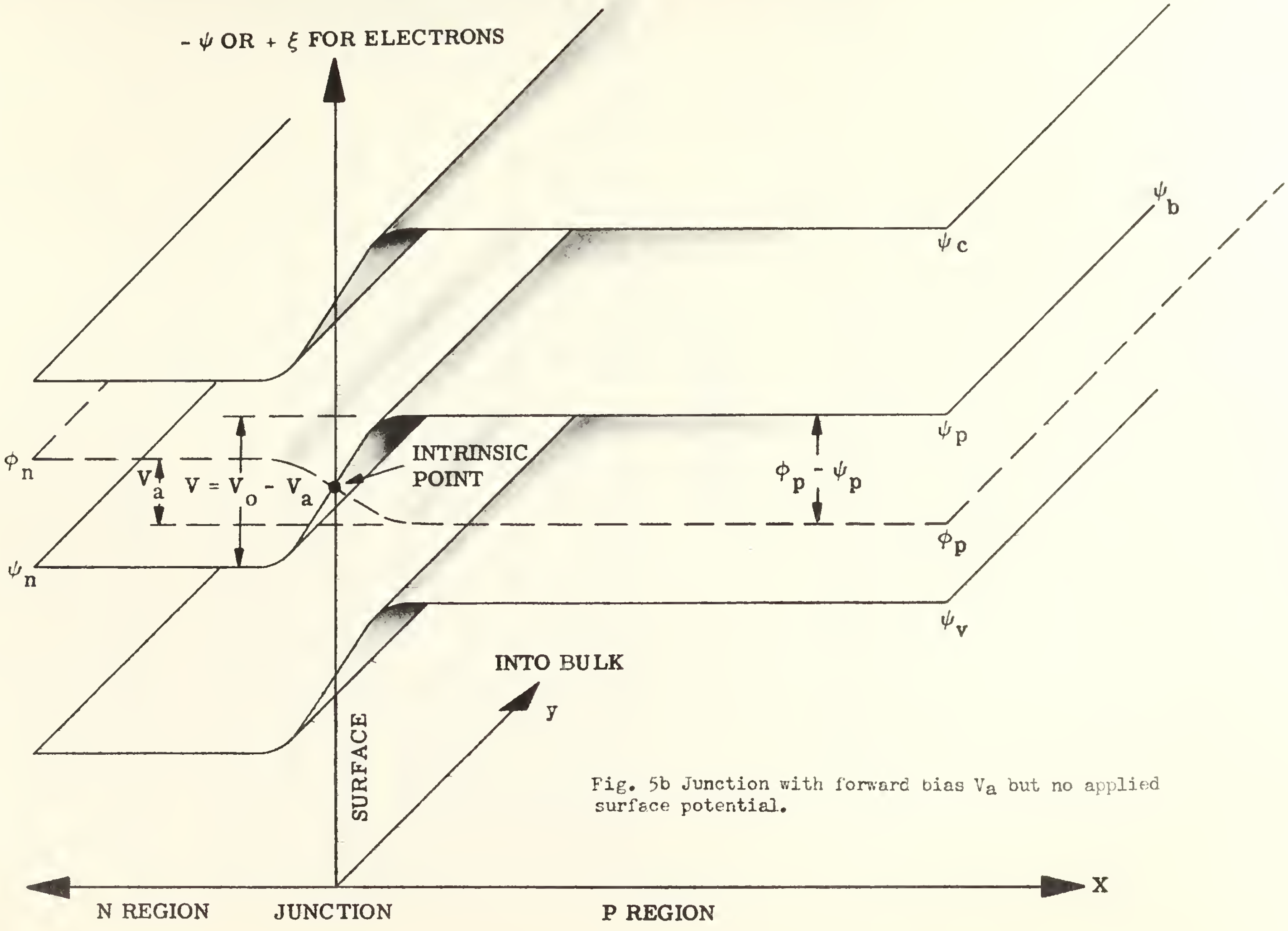


Fig. 5b Junction with forward bias V_a but no applied surface potential.

potential barrier for electrons which allows the higher energy electrons on the n side to cross to the p side. This is conventional diode operation. It is usually stated as the change in the quasi Fermi levels. This is also shown in figure 5(b). The concentration of holes in general may be represented by the quantity $\varphi - \psi$. This relationship is:

$$p \doteq n_i e^{q(\varphi-\psi)/kT} \tag{1}$$

where n_i is the intrinsic charge carrier concentration. It is a function of temperature only for a given ψ. The corresponding expression for electron concentration is:

$$n \doteq n_i e^{-q(\varphi-\psi)/kT} \tag{2}$$

It is seen that the quantity $\varphi-\psi$ is positive on the p side and hence $p_p \gg n_p$ as expected (the ordinate of figure 5 is $-\psi$) while $\varphi-\psi$ is negative on the n side and $n_n \gg p_n$. It should be pointed out that applying the bias, V_a, did not materially change the quantity $\varphi-\psi$ and so the carrier concentrations have not been modified. This is only true where the injected carrier density is small compared with the equilibrium majority concentration on the other side of the barrier. Therefore we may avoid explicit mention of bias conditions and discuss the effects of bias in terms of the quantity $\varphi-\psi$.

The effect of applying a positive potential to the grid is shown in figure 5(c). As previously assumed, it has no effect upon the n side. The effect upon the p side is to lower the energy of all of the allowed states in the region to which the surface-applied electric field penetrates. There are now more electrons with energies greater than or equal to the new energy of these previously empty states. These electrons will make transitions to these empty states which raises the Fermi level toward

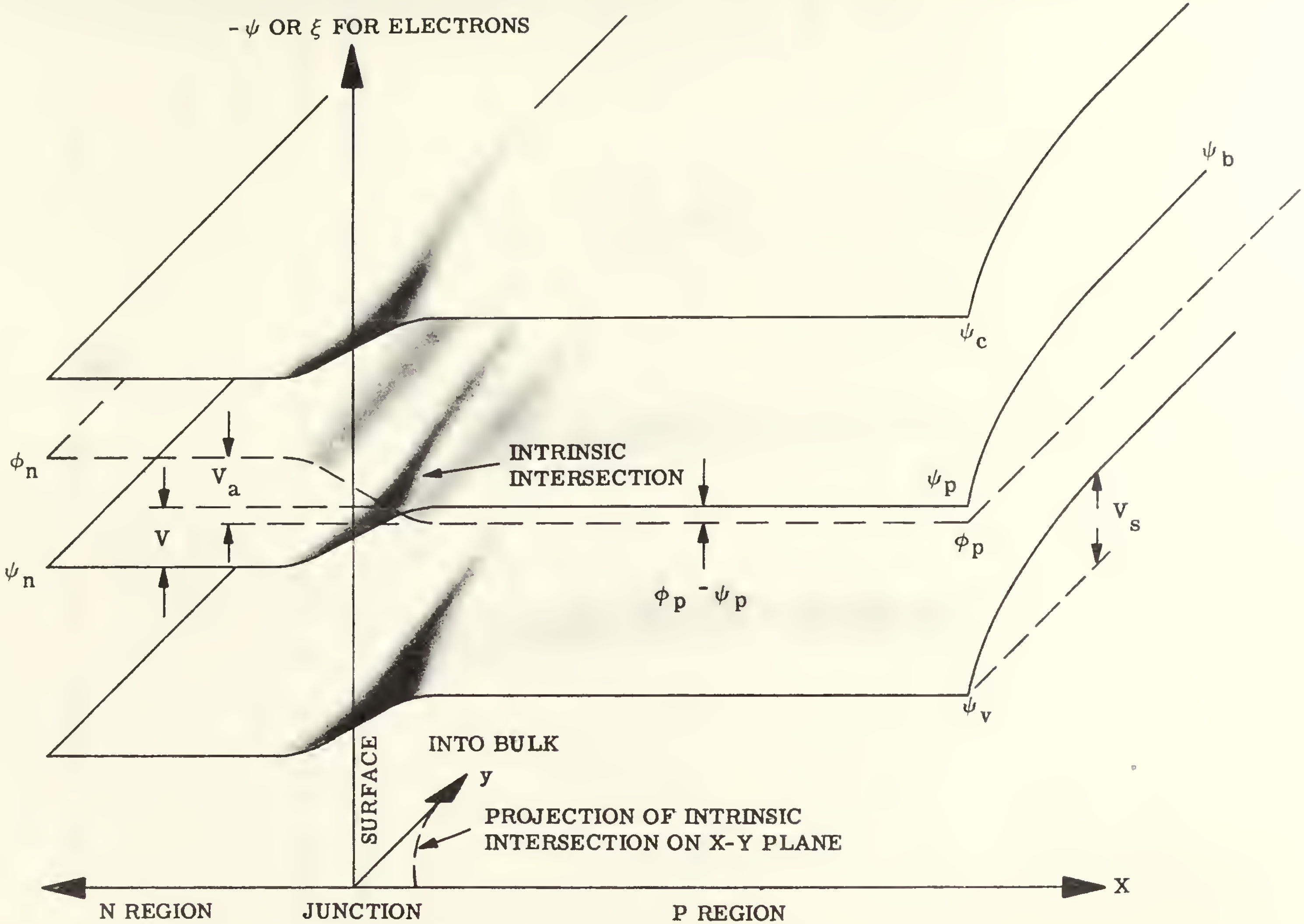


Fig. 5c Junction with forward bias V_a and positive potential V_s applied to the surface.

the conduction band. (The energy at which the probability of an electron occupying a state at that energy is one half, has been moved toward the conduction band. In figure 5(c), φ_p is the reference level so the moving of the Fermi level is manifested by the bending of the energy levels down toward the Fermi level.) In terms of the behavior of the semiconductor at the surface, the quantity $\varphi - \psi$ reflects the change. $\varphi - \psi$ has become much smaller than its previous value on the p side. As a result, the exponent of the expression for the concentration of holes is smaller and therefore p is smaller. Conversely the exponent in the expression for n increases and n is larger. The two are more nearly equal and the material more nearly intrinsic.

An interesting point to note in passing is that the conductivity of the surface layer increases as the concentration of electrons increases. This is due to the higher mobility of electrons over that of holes:

$$\sigma = q(\mu_n n + \mu_p p)$$

$$= q\mu_n \left(n + \frac{\mu_p}{\mu_n} p\right)$$

where $\mu_p/\mu_n < 1$. σ is the semiconductor conductivity, μ_n and μ_p are the respective mobilities for electrons and holes. The conductivity is seen to increase with n. Since ψ decreases as a function of depth into the material, so does n and consequently σ. Therefore the voltage drop due to a longitudinal channel current is less near the surface of the channel than deeper into the channel. This is another factor which tends to confine channel current to the immediate surface layer.

Increasing the grid voltage further results in the situation shown in figure 5(d). The energy bands have been lowered so far that the Fermi level is now between ψ and the conduction band. The sign of $\varphi - \psi$ has changed and n is now greater than p. The surface layer is now n-type

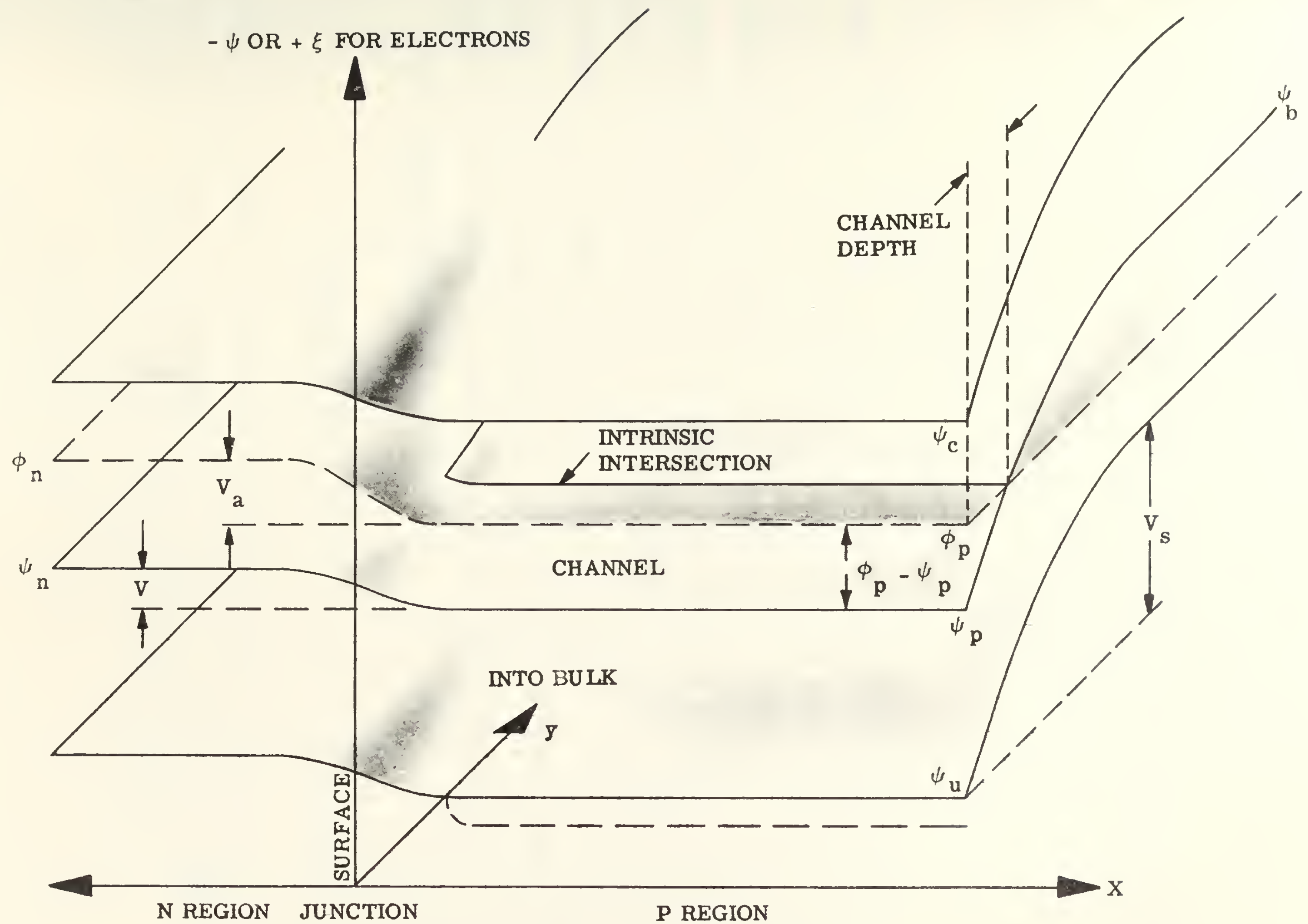


Fig. 5d Junction with forward bias V_a and applied surface potential V_s. Notice that the intrinsic intersection no longer touches surface. An n-type layer has been formed on the surface.

material. This satisfies the previous definition of channel. As can be seen from figure 5(d), the n layer extends into the p-type material to a depth determined by the intersection of ψ and φ .

A determination of the depth of the channel may be made by solving Poisson's equation for the potential in the interior of the p region as a function of the grid-induced surface potential:

$$\nabla^2 \psi(r) = -\rho(r)/\epsilon$$

where $\rho(r)$ is the charge density as a function of position and ϵ is the permittivity of the semiconductor. The approach will parallel the analysis presented by Kingston and Neustadter /8/. It will be assumed that the variation of potential is unidirectional and not a function of time. The surface is the origin of the y coordinate of figure 5. The energy origin is taken as the Fermi level in the p-type material. The variation of potential in the transition region is neglected so that our discussion is limited to the region in which the grid has the greatest effect.

The charge density is given by:

$$\rho = N_d - N_a + p - n$$

Deep in the bulk it is known that charge neutrality exists and the electric field is zero. The potential has some equilibrium value ψ_B shown in figure 5. Using the charge neutrality condition we have:

$$\rho = N_d - N_a + p - n = 0$$
$$n - p = N_d - N_a$$

n and p are given by (1) and (2)

$$(n-p) = n_i \left\{ e^{q(\psi-\varphi)/kT} - e^{-q(\psi-\varphi)/kT} \right\}$$
$$= 2n_i \sinh q\psi/kT$$

Deep in the bulk $\psi = \psi_B$ and $n-p = N_d - N_a$, hence:

$$N_d - N_a = 2\,n_i \sinh q\psi_B/kT$$

Similarly $p-n = -2n_i \sinh q\psi/kT$ in general.

Poisson's equation for the one dimensional case becomes

$$\frac{d^2\psi}{dx^2} = \frac{2n_i}{\epsilon}\left[\sinh q\psi/kT - \sinh q\psi_B/kT\right] \tag{3}$$

Using the identity

$$\frac{d^2\psi}{dx^2} \quad \frac{d}{d\psi}\left(\frac{d\psi}{dx}\right)^2$$

and integrating, (3) becomes:

$$\int_0^{\frac{d\psi}{dx}} d\left(\frac{d\psi}{dx}\right)^2 = \int_0^{\psi} \frac{2n_i}{\epsilon}\left[\sinh q\psi/kT - \sinh q\psi_B/kT\right]d\psi$$

This integration is from the bulk toward the surface and yields an expression for the electric field as a function of the potential at any point:

$$\frac{d\psi}{dx} = \sqrt{\frac{2n_i}{\epsilon}}\left[\frac{kT}{q}\cosh\frac{q\psi}{kT} - \frac{kT}{q}\cosh\frac{q\psi_B}{kT} - (\psi-\psi_B)\sinh\frac{q\psi_B}{kT}\right]^{\frac{1}{2}} \tag{4}$$

Equation (4) must now be integrated to find the potential ψ as a function of x:

$$\int_0^x dx = \int_{\psi_S}^{\psi} \frac{d\psi}{\sqrt{\frac{2n_i}{\epsilon}}\left[\frac{kT}{q}\left(\cosh\frac{q\psi}{kT} - \cosh\frac{q\psi_B}{kT}\right) - (\psi-\psi_B)\sinh\frac{q\psi_B}{kT}\right]^{\frac{1}{2}}} \tag{5}$$

The solution of this integral will yield the potential at any depth. More importantly, if $\psi = 0$ is the upper limit of the right hand integral, the corresponding value of x will be a measure of the depth of penetration of the channel into the bulk as a function of ψ_S. A solution in closed form is not possible however a plausibility argument will show

that the channel depth saturates as a function of ψ_s .

The initial value of ψ_s is positive and ψ_B is negative. The integrand is the slope of the integral. For large positive values of ψ_s the denominator approaches infinity with the cosh $\frac{q\psi}{kT}$ hence the integrand approaches zero. For smaller values of ψ_s the slope will have a measurable magnitude. For fairly large ψ_s $(> 5\frac{kT}{q})$ the integrand will be essentially $\exp(-q\psi_s/kT)$ which rapidly approaches zero as an asymptote. Thus it can be seen that increasing ψ_s will not materially affect the integral and hence the depth of the channel.

4. The Phenomenon of Recombination in Semiconductors

The phenomenon of recombination of electrons and holes in semiconductors has been observed largely in connection with photoconduction processes. In such a process, hole-electron pairs are generated by absorption of energy from light incident upon a semiconductor. With an external electric field applied these carriers constitute a drift current which is observed to decay exponentially with distance from the incident light source. Such decay can only be explained by a reduction in the number of charge carriers and has been described as a recombination of the light generated hole-electron pairs. This is a random process which is best described statistically. The most widely accepted model for recombination processes was described by W. Shockley and W.T. Read, Jr./10/ Their model was concerned with the recombination rate in a bulk material with no p-n junctions. This model was used by C.T. Sah, R.N. Noyce and W. Shockley to describe the same non-saturable leakage current and low forward current characteristics in silicon p-n junctions as were explained with channel considerations by Cutler and Bath.

Recombination may be described as either one or two step processes. The one step process is a direct encounter of an electron by a hole which results in the removal of both from the conduction process. Essentially the electron has made a transition from the conduction band to fill a vacancy in the valance bond structure. This process is unlikely since the electron must make an energy-losing collision with some other particle at the precise point in time and space where the hole is located. Since the hole is also moving through the crystal this is tantamount to a man throwing one baseball randomly into the air and another to a batter. If the second baseball makes a collision with the bat and subsequently hits the first ball you have described a direct collision of an electron

and a hole.

The two step process involves an intermediate "trapping" of either the electron or the hole in a localized volume in space. Since its probability of being at a given point in space is now greater, the probability of being encountered or hit by another particle is greater. The effect of traps is to increase the rate of recombination over the case where no traps are present.

Any localized imperfection in the crystal may act as a trap. An imperfection in the crystal may take the form of a vacancy in the lattice, an impurity atom, a crystal dislocation, an interstitial atom, or a grain boundary. Since the surface of a semiconductor always defines a crystal grain boundary, there will always be a certain concentration of traps at the surface. This is reflected in the relative importance of surface recombination as opposed to recombination in the bulk material. The presence of a crystal imperfection is manifested by an allowed state within the forbidden band between the valence and conduction bands. The potential level of this allowed state is dependent upon the type and location of the imperfection. Another way of describing a trap is in terms of a localized electric field which attracts an electron or hole and holds it in a local area whose radius is dependent upon the strength of the local field. If the field attracts an electron very strongly, it corresponds to nearly immobilizing the electron which makes it very easy for a hole to subsequently capture it. The corresponding energy level of the trap would be near the valence band.

The rate at which recombination takes place describes a loss of current. The fact that recombination also takes place in channels establishes an inter-relation between the two phenomena which is important to the description of transistor action in the tetrode. The following

analytical description will parallel Shockley and Read /10/ and Sah, Noyce and Shockley /11/, arriving at an expression given by Sah for the recombination rate as a function of the electrostatic potential at the surface and in the bulk. Using this expression we may then correlate the recombination current in the channel with the depth of the channel and surface potential. This should allow us to separate the effects of surface potential upon channel depth and recombination rate.

Let us suppose that in our material there is a density of traps N_t. It will be assumed that these traps all have the same energy level. This assumption is relatively valid for a well purified semiconductor which has been doped while carefully avoiding unwanted impurities and careful crystal growing to eliminate lattice vacancies and dislocations.

A trap may have one of two conditions: it may be occupied by an electron or not occupied by an electron. It is to be remembered that a trap is simply a quantum energy state or "allowed energy" for an electron. As such its probability of being occupied by an electron is described by the Fermi-Dirac statistics in exactly the same manner as electrons and holes. If a trap is occupied by an electron, it may either emit that electron to the conduction band or capture a hole from the valence band. (The electron actually makes a second transition to the valence band to fill a vacant bond.) In either case the end result is that the trap is empty. If a trap is initially empty, it may capture an electron from the conduction band or emit a hole to the valence band. (In the latter case a valence bond is broken and the released electron is trapped rather than going to the conduction band.) The end result in either case is that the empty state is filled by an electron. Obviously the density of filled traps plus the density of empty traps is equal to the total density of traps. Thus the probability that a trap is full plus the probability

that it is empty is unity.

The probability that any allowed quantum state is filled is given by the Fermi-Dirac distribution:

$$f(E) = \frac{1}{1 + e^{(E-F)/kT}}$$

where E is the energy of the state and F, known as the Fermi level, is the energy of a state which would have an equal probability of being filled or empty. Since kT is of the order of 4×10^{-21} electron volts, $e^{(E-F)/kT}$ is usually much larger than 1 and f(E) can be approximated with almost no error by:

$$f(E) \doteq e^{-(E-F)/kT} = e^{(F-E)/kT}$$

Since $E = -q\psi$ where ψ is potential (or voltage):

$$f(E) \doteq e^{-q(\varphi-\psi)/k} = e^{q(\psi-\varphi)/kT}$$

The probability of a state being empty is of course:

$$f_p(E) = 1 - f(E) = 1 - \frac{1}{1+e^{(E-F)/kT}} = \frac{e^{(E-F)/kT}}{1+e^{(E-F)/kT}}$$

This can also be written as:

$$f_p(E) = f(E)\, e^{(E-F)/kT} = \frac{1}{1+e^{-(E-F)/kT}}$$

The probability that a trap is full must then be given by:

$$f_t = f(E_t) \doteq e^{(F_t-E_t)/kT}$$ and empty by

$$f_{pt} = f_p(E_t) = 1 - f_t$$

where E_t is the energy level of the trap and F_t is the Fermi level for the traps. We also define F_n and F_p as being the Fermi levels for electrons and holes respectively. Under conditions of thermal equilibrium (no applied biases) all of these Fermi levels coincide. However, when the concentrations of carriers are not at their equilibrium values, the

amount by which they differ is reflected in a change in Fermi level.

Our objective is to determine the rate of recombination of electrons and holes. In order to do this we shall determine the rate at which electrons are being trapped by the empty traps and the rate at which full traps are emitting their electrons to the conduction band. The difference will be the net electron capture rate. Similarly, the rate of hole capture and emission will be found in order to determine the net hole capture rate. At steady-state (but not equilibrium) conditions, the net hole capture rate and net electron capture rate must be equal. That is, the rate of filling empty traps must equal the rate of emptying filled traps. This, then, is the rate at which carriers are being removed from the conduction process and is the rate of recombination. For every hole-electron pair which recombines, one electron must flow in the external circuit as compensation. This constitutes a recombination current. We will have more to say about this later. The energy which is given up when an electron and a hole recombine is converted to lattice vibrational energy. This must be considered as part of the total power dissipation of the tetrode. More will be said about this in Section 9 when behavior of the tetrode will be discussed.

The rate at which electrons are captured is dependent upon; (1) the number of electrons available in the conduction band for capture; (2) the number of traps which are empty and hence capable of trapping an electron; and (3) the probability per unit time that an electron with a given energy will be trapped by an empty trap at a given energy. The product of these factors yields the rate of electron capture. Analogously, the rate of electron emission from full traps to the conduction band is given by the product of (1) the number of empty states in the conduction band as a function of energy; (2) the number of full traps (hence capable of emitting

an electron); and (3) the probability per unit time of a full trap emitting an electron to a given energy level in the conduction band.

The number of electrons available in the conduction band, n, is given by the product of the number of states and the probability of that state being occupied integrated over the conduction band. For a non-degenerate* semiconductor, the effective density of states in the conduction band acts like a number of states, N_c, all concentrated at the bottom of the conduction band. Hence the electron density integral degenerates to:

$$n = \int_{E_c}^{\infty} N(E)\, f(E)\, dE = N_c f(E_c) = N_c\, e^{(F_n - E_c)/kT} \qquad (5)$$

$$N(E) = \begin{cases} 0; & E \neq E_c \\ N_c; & E = E_c \end{cases}$$

The number of traps which are empty is given by the product of the density of traps and the probability of a trap being empty integrated over all energies. It has been assumed that a density of traps, N_t, is concentrated at the single energy E_t. The integral degenerates as in (5) above:

$$n_{pt} = \int_0^{\infty} N_t\, f_{pt}(E) = N_t f_{pt}(E_t) = N_t f_t\, e^{(E_t - F_t)/kT}$$

The probability per unit time that an electron with a given energy will be trapped is given by the product of the thermal velocity of the electron and the capture cross section. This is a constant with respect to energy for a non-degenerate semiconductor since all of the conduction band electrons are effectively at E_c. This factor shall be called c_n.

*A non-degenerate semiconductor is one which is sufficiently pure that the electrons in the conduction band are essentially at the lowest end of the band. This implies that the density is still a function of temperature. If the density of electrons becomes too high the concentration loses its temperature dependence and the semiconductor takes on the properties of a metal. This latter condition is known as a degeneracy.

Using these results, the rate of electron capture is:

$$n\, n_{pt}\, c_n = N_c\, f(E_c)\, N_t f_{pt}(E_t)\, c_n \qquad (6)$$

Considering the rate of emission in an analogous manner, the number of empty states in the conduction band is $N_c - n$ which is:

$$N_c - n = N_c\, f_p(E_c) = N_c\, f(E_c)\, e^{(E_c - F_n)/kT}$$

The number of full traps is:

$$n_t = N_t\, f_t(E_t)$$

and the corresponding probability of emission per unit time is e_n. These factors give the rate of electron emission to the conduction band as:

$$(N_c - n)\, n_t\, e_n = N_c\, f_p(E_c)\, N_t f_t(E_t)\, e_n \qquad (7)$$

The net rate of electron capture is found by differencing (6) and (7).

$$\begin{aligned} U_{cn} &= N_c f(E_c) N_t f_{pt}(E_t) c_n - N_c f_p(E_c) N_t f_t(E_t) e_n \\ &= N_c N_t c_n \left[f(E_c) f_{pt}(E_t) - f_p(E_c) f_t(E_t) \frac{e_n}{c_n} \right] \end{aligned}$$

If the thermal equilibrium argument is invoked, the rate of emission and the rate of capture must be equal and $F_t = F_n$. Under this condition $U_{cn} = 0$ which means:

$$f(E_c) f_{pt}(E_t) = f_p(E_c) f_t(E_t) \frac{e_n}{c_n}$$

$$f(E_c) f_t(E_t)\, e^{(E_t - F_t)/kT} = f(E_c)\, e^{(E_c - F_n)/kT} f_t(E_t) \frac{e_n}{c_n}$$

$$e^{(E_t - E_c)/kT} = \frac{e_n}{c_n} \qquad (8)$$

Since E_t is always less than E_c, the rate at which electrons are emitted from the traps <u>to the conduction band</u> under equilibrium conditions is always less than the rate at which electrons are falling into the traps <u>from the conduction band</u>. This should not be construed to mean that the traps can never be emptying as fast as they are filling since we have not yet considered the other means of emptying traps: hole capture.

Returning to the expression for U_{cn} and substituting (8)

$$U_{cn} = N_c f(E_c) N_t c_n f_{pt}(E_t)\left[1 - e^{(E_t-E_c)/kT} e^{(E_c-F_n)/kT} e^{(F_t-E_t)/kT}\right]$$
$$= n f_{pt}(E_t) C_n \left[1 - e^{(F_t-F_n)/kT}\right]$$

Expanding this expression and isolating the exponential part:

$$n f_{pt}(E_t) C_n e^{(F_t-F_n)/kT} = N_c e^{(F_n-E_c)/kT} f_t e^{(E_t-F_t)/kT} C_n e^{(F_t-F_n)/kT}$$
$$= N_c e^{(E_t-E_c)/kT} C_n f_t$$

It will be found convenient to express the electron emission term in terms of a ficticious electron density which has all of the characteristics of an electron distribution with the Fermi level at E_t. Designate this by

$$n_1 \equiv N_c e^{(E_t-E_c)/kT} \quad \text{then} \tag{9}$$
$$U_{cn} = n f_{pt}(E_t) C_n - n_1 f_t(E_t) C_n \tag{10}$$

The term $C_n = N_t c_n$ expresses the number of electrons per unit time which would be captured if all of the traps were empty. The reader should not try to attach meaning to the individual terms in the emission rate term of (10). The physical meanings have been distorted by the mathematical manipulation. The only significant thing to remember is that n_1 describes mathematically an electron concentration analogous to n but with the Fermi level at E_t. This will be found useful.

Turning now to the net capture rate for holes, the same procedure may be followed in obtaining U_{cp}. The rate of hole capture is the product of hole density, filled traps and a capture rate term, c_p. The rate of hole emission is the product of the number of filled holes in the valence band, N_v-p, empty traps and an emission rate term. The net hole capture rate is:

$$U_{cp} = p N_t f_t(E_t) c_p - (N_v - p) f_{pt}(E_t) e_p$$

$$U_{cp} = N_v N_t c_p \left[f_p(E_v) f_t(E_t) - f(E_v) f_{pt}(E_t) \frac{e_p}{c_p} \right]$$

Equilibrium arguments determine e_p/c_p to be $e^{(E_v-E_t)/kT}$. If p_1 is defined as

$$p_1 \equiv N_v e^{(E_v-E_t)/kT} \qquad \text{then} \qquad (11)$$

$$U_{cp} = p f_t(E_t) C_p - p_1 f_{pt}(E_t) C_p \qquad (12)$$

where C_p is defined as $N_t c_p$ analogously to C_n.

If steady state conditions are assumed, the net rates of hole and electron capture are equal and this rate becomes the recombination rate. Equating U_{cp} and U_{cn} and solving the equation first for f_{pt} and then for f_t:

$$p f_t(E_t) C_p - p_1 f_{pt}(E_t) C_p = n f_{pt}(E_t) C_n - n_1 f_t(E_t) C_n$$

$$p\left\{1 - f_{pt}(E_t)\right\} C_p - p_1 f_{pt}(E_t) C_p = n f_{pt}(E_t) C_n - n_1 \left\{1 - f_{pt}(E_t)\right\} C_n$$

$$f_{pt} = \frac{p C_p + n_1 C_n}{C_p(p+p_1) + C_n(n+n_1)}$$

The substitution of $f_{pt} = 1 - f_t$ yields:

$$f_t = \frac{n C_n + p_1 C_p}{C_p(p+p_1) + C_n(n+n_1)}$$

When these expressions are substituted back into (10) or (12) it is found that:

$$U = \frac{pn C_n C_p + p p_1 C_p^2 - p_1 p C_p^2 - p_1 n_1 C_n C_p}{C_p(p+p_1) + C_n(n+n_1)}$$

$$= \frac{pn - p_1 n_1}{(p+p_1)/C_n + (n+n_1)/C_p} \qquad (13)$$

If C_p and C_n, which have the dimensions of reciprocal seconds, are defined as the lifetimes of holes in highly n-type material and electrons in highly p-type material respectively,

$$\tau_{po} \equiv 1/C_p \quad ; \quad \tau_{no} \equiv 1/C_n \qquad \text{then}$$

$$U = \frac{pn - p_1 n_1}{(p+p_1)\tau_{no} + (n+n_1)\tau_{po}} \qquad (14)$$

In the argument so far, equilibrium conditions have not been invoked except to evaluate constants. The expression for recombination rate (13) is valid for steady state non-equilibrium conditions. If the product $p_1 n_1$ is evaluated using (9) and (11) it is found to be

$$p_1 n_1 = N_v e^{(E_v - E_t)/kT} \; N_c e^{(E_t - E_c)/kT} = N_c N_v e^{(E_v - E_c)/kT} \equiv n_i^2$$

It is now convenient to recall that the ficticious carrier concentrations were described in the same manner as the true carrier concentrations. Also recall that the carrier concentrations can be expressed in terms of an intrinsic energy level, an intrinsic concentration and the Fermi level. The reference for energy has been established as the Fermi level so the usual method of expressing variation of carrier concentrations in terms of quasi Fermi levels for electrons and holes will not be used. Instead these variations will be expressed in terms of the variable ψ. The Fermi level for holes in the p-type bulk will be established as the reference. The same approach may be used for n_1 and p_1.

$$n = N_c e^{(F_n - E_c)/kT} = n_i e^{(F_n - E_i)/kT} = n_i e^{q\psi/kT}$$

$$n_1 = N_c e^{(E_t - E_c)/kT} = n_i e^{(E_t - E_i)/kT} = n_i e^{q(\psi - \psi_t)/kT}$$

$$p = N_v e^{(E_v - F_p)/kT} = n_i e^{(E_i - F_p)/kT} = n_i e^{q\psi_B/kT}$$

$$p_1 = N_v e^{(E_v - E_t)/kT} = n_i e^{(E_i - E_t)/kT} = n_i e^{q(\psi_t - \psi)/kT}$$

These quantities substituted in (14) give an expression for recombination of the form derived by Sah and Shockley:

$$U = \frac{pn - n_i^2}{(p + p_1)\tau_{no} + (n + n_1)\tau_{po}} = \frac{n_i^2 \left[e^{q(\psi - \psi_B)/kT} - 1 \right]}{n_i \left[\tau_{no}\left(e^{-q\psi_B/kT} + e^{q(\psi_t - \psi)/kT} \right) + \tau_{po}\left(e^{q(\psi - \psi_t)/kT} + e^{q\psi/kT} \right) \right]}$$

$$= \frac{n_i}{\sqrt{\tau_{no}\tau_{po}}} \left[\frac{e^{q(\psi - \psi_B)/2kT} \left(e^{q(\psi - \psi_B)/2kT} - e^{-q(\psi - \psi_B)/2kT} \right)}{e^{-q\psi_B/kT - \ln\sqrt{\frac{\tau_{po}}{\tau_{no}}}} + e^{q(\psi_t - \psi)/kT - \ln\sqrt{\frac{\tau_{po}}{\tau_{no}}}} + e^{-q(\psi_t - \psi)/kT + \ln\sqrt{\frac{\tau_{po}}{\tau_{no}}}} + e^{q\psi/kT + \ln\sqrt{\frac{\tau_{po}}{\tau_{no}}}}} \right]$$

$$U = \frac{n_i}{\sqrt{\tau_{no}\tau_{po}}}\left[\frac{\sinh\left\{\frac{q(\psi-\psi_B)}{2kT}\right\}}{\cosh\left\{\frac{q(\psi+\psi_B)}{2kT}+\ln\sqrt{\frac{\tau_{po}}{\tau_{no}}}\right\}+e^{q(\psi_B-\psi)/2kT}\cosh\left\{\frac{q(\psi_t-\psi)}{kT}-\ln\sqrt{\frac{\tau_{po}}{\tau_{no}}}\right\}}\right] \tag{15}$$

This expression now gives the rate of recombination of electrons and holes in the channel and the junction transition region as a function of the intrinsic Fermi level. This expression may be used to determine the current in the channel and the transition region. A correlation of this result with the depth of the channel should reveal the connection if any between channel current and depth of channel.

5. Effect of Channel and Recombination on Transistor Action

In sections 3 and 4 the models for channel and recombination were developed and their dependence on the surface potential was given. Now it is necessary to describe what effect these two phenomena have on the device as a transistor.

Consider a transistor biased in the usual manner. The emitter-base junction is forward biased and the collector-base junction is reverse biased. For an n-p-n transistor, electrons are the minority carriers in the base and make up most of the current which flows. Holes from the base constitute a small fraction of the total however. The ratio of electron current to total current in the emitter base junction is given by the emitter efficiency. For a given junction voltage, a total current given by the usual diode equation will flow.

Now apply a potential to the grid of the tetrode sufficient to create a channel such as described in section 3. Electrons will flow from the emitter into the channel. However this in itself would not be any cause for concern for it does not constitute a current in the steady state sense. The electrons have not crossed a junction into the base because the junction effectively lies between the channel and the base. But recombination of an electron and a hole will occur in the channel at a rate depending on the surface potential and given by equation (15) of section 4. This is at the heart of the tetrode's functioning. For every electron-hole recombination, an electron must flow into the channel from the emitter and a hole must enter the channel from the base to restore the channel to its former state. This is again determined by the surface potential as was shown in section 3 by equation (5). The electron which comes from the emitter does not cross a junction and hence does not constitute current in the sense of the diode equation. However the hole

coming from the base does cross the junction and hence does carry current. The number of recombinations which takes place depends upon the channel volume and the rate per unit volume at which these recombinations are taking place. As was shown in section 3, the size of the channel becomes nearly a constant very quickly so the major effect is that of the changing recombination rate with changing surface potential.

It was previously stated that the ratio of electron current to total current was the emitter efficiency. The total current is fixed by the junction voltage. Thus for every hole which crosses the junction into the channel, one less electron crosses the emitter junction into the base where it subsequently diffuses to the collector and becomes collector current. It is fairly apparent that the more holes which flow into the channel, the fewer electrons will flow into the base and from there to the collector. Indeed it is possible for the entire current required by the diode equation to be made up of holes flowing into the channel thus reducing electron current (and thereby collector current) to zero. This can also be described in terms of the emitter efficiency being reduced to zero.

The current transfer ratio, α, is the ratio of current crossing the emitter-base junction to the current crossing the collector base junction. It is usually considered to be made up of three factors; (1) γ, the emitter efficiency; (2) β, the transport factor; and (3) μ, the collector multiplication factor. μ is normally unity for the usual transistor bias conditions. β describes the number of minority carriers which cross the emitter-base junction and the collector-base junction. It is a measure of the minority carrier recombination in the base bulk. Modern technology has made this factor nearly unity also. This leaves α as being for the most part dependent on γ. In modern transistors this is almost

urity also which accounts for the very high values of α encountered. It is true that:

$$\beta = \frac{\alpha}{1-\alpha}$$

where β is now the common emitter current transfer ratio, not the transport factor. It can be seen that as recombination in the channel degrades the emitter efficiency, α is reduced and consequently β is reduced quite rapidly. This will be seen to be the case in section 7 where experimental curves show the effect of grid voltage on β.

6. Hybrid Equations and Parameters

The tetrode is a four terminal device hence, relative to one of these terminals, it may be described in terms of the current and voltage at each of the remaining three terminals. A simple extension of the hybrid equations for a transistor adequately describes the tetrode. In terms of the notation of figure 6, the hybrid equations are:

$$V_1 = h_{11}I_1 + h_{12}V_2 + h_{13}V_3$$

$$I_2 = h_{21}I_1 + h_{22}V_2 + h_{23}V_3$$

$$I_3 = h_{31}I_1 + h_{32}V_2 + h_{33}V_3$$

The defining expressions for the h-parameters are the same as for the conventional transistor parameters. Note that h_{11}, h_{12}, h_{21} and h_{22} are virtually the same parameters as in the transistor case, the only difference being that it is also necessary to specify the short circuited condition of the grid (terminal 3). The equivalent circuit for the tetrode will be developed from a physical interpretation of the parameters. The common emitter configuration will be used.

The voltage at the base terminal is made up of three terms reflecting the conditions of the base, grid and collector.

$z_B = h_{11} = \left.\frac{\partial V_1}{\partial I_1}\right|_{V_2=V_3=0}$ is the familiar expression for the base input impedance and is described with the collector and grid short circuited.

$\mu_{CB} = h_{12} = \left.\frac{\partial V_1}{\partial V_2}\right|_{I_1=V_3=0}$ is a voltage feedback ratio reflecting the effect of collector voltage on base voltage. The base is open and the grid short-circuited. This parameter is manifested in the equivalent circuit by a voltage controlled voltage source.

$\mu_{GB} = h_{13} = \left.\frac{\partial V_1}{\partial V_3}\right|_{I=V_2=0}$ is also a voltage feedback ratio reflecting the effect of grid voltage on base drive. The base is open and the collector short-circuited. This parameter is

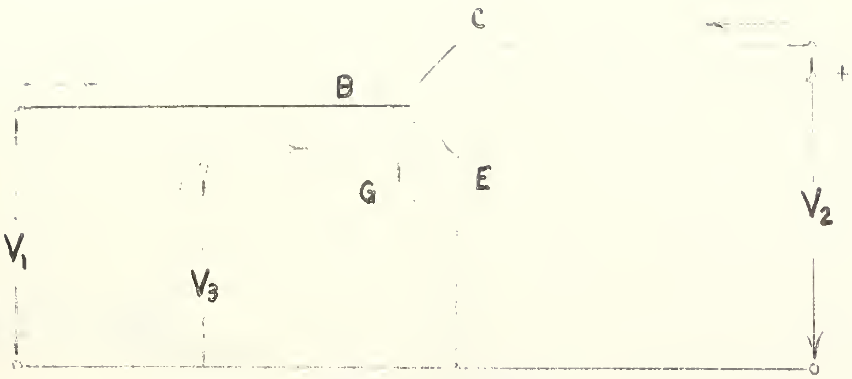


Fig. 6 Tetrode symbol and current-voltage conventions

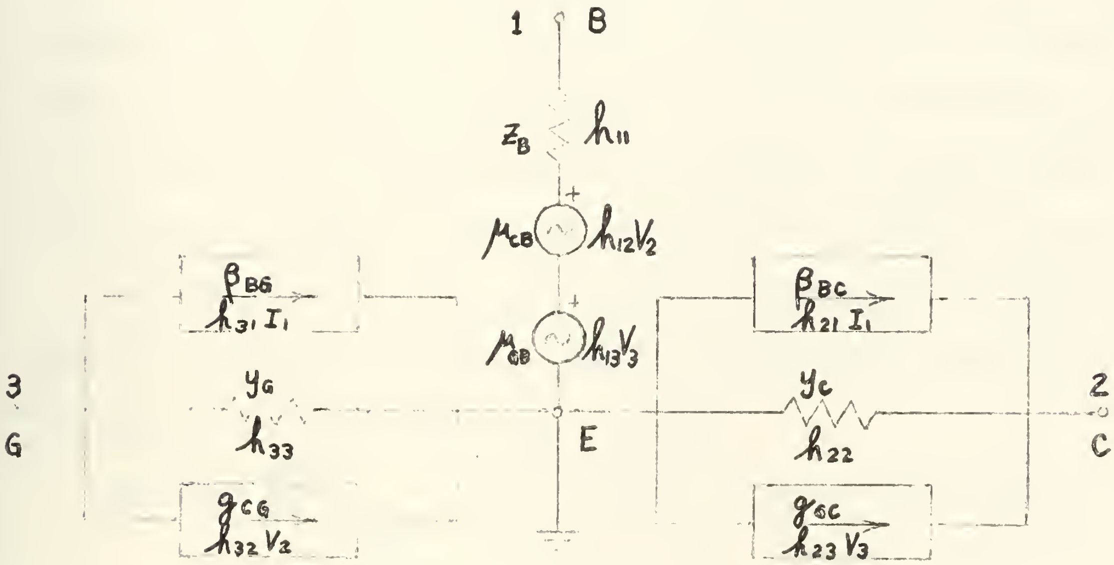


Fig. 7 Tetrode linear equivalent circuit

also represented by a voltage controlled voltage source.

The three terms combine in series to form branch 1 of figure 7. The collector current is to be made up of three terms which will add in parallel to form branch two.

$\beta_{BC} = h_{21} = \left.\dfrac{\partial I_2}{\partial I_1}\right|_{V_2=V_3=0}$ is the forward base-collector short circuit current transfer ratio. It is defined with the collector and grid short-circuited. In the equivalent circuit it is represented by a current controlled current generator.

$y_c = h_{22} = \left.\dfrac{\partial I_2}{\partial V_2}\right|_{I_1=V_3=0}$ is the collector admittance since it has the dimensions of mhos. The base is open and the grid short-circuited for this measurement.

$g_{GC} = h_{23} = \left.\dfrac{\partial I_2}{\partial V_3}\right|_{I_1=V_2=0}$ is perhaps the most significant parameter of the tetrode. It describes the effect upon collector current caused by a potential impressed upon the high impedance grid. It has the dimensions of a conductance and describes the grid-to-collector transfer effect. Hence it is the grid-collector transconductance, entirely analogous to a vacuum tube grid-plate transconductance. It is defined with the base open and the collector short-circuited. In the equivalent circuit it is shown as a voltage controlled current source.

The third branch of the hybrid equivalent circuit is also descriptive of a current and so will be made up of equivalent shunt elements. It describes the components of grid current.

$\beta_{BG} = h_{31} = \left.\dfrac{\partial I_3}{\partial I_1}\right|_{V_2=V_3=0}$ is a base-grid current transfer ratio of the same type as h_{21}. However, it describes a feedback factor rather than a forward transfer ratio. It takes the form of a current controlled current generator in the equivalent circuit.

$g_{CG} = h_{32} = \left.\dfrac{\partial I_3}{\partial V_2}\right|_{I_1=V_3=0}$ is another feedback conductance giving the effect of collector voltage on grid current.

It takes the form of a voltage controlled current generator in the equivalent circuit.

$$y_3 = h_{33} = \left.\frac{\partial I_3}{\partial V_3}\right|_{I_1 = V_2 = 0}$$

is the grid input admittance. It should consist of a pure capacity for all intents and purposes since the conductance is very nearly zero through the grid oxide layer.

This then describes the tetrode in terms of measurable parameters provided the driving signal is small. It remains to describe these parameters and their variations with bias conditions. The behavior of the tetrode as a transducer will then be predictable for circuit applications.

F. [illegible] of [illegible] Variations in Terms of the [illegible]

In this section a profile of each parameter will be presented to demonstrate the manner in which the hybrid small-signal parameters vary with bias. Unfortunately, owing to transistor failure and the difficulty in obtaining replacements in a short space of time, a complete set of data for all bias parameters was not achieved. Data was not obtained to describe Y_c, S_{cg} and β_{bg}.

When thinking about the parameters and their variations, it is well to keep in mind that the device is capable of either a current mode or voltage mode of operation as a function of the grid bias voltage. For positive grid voltages the tetrode operates best in the voltage mode with a signal applied to the grid. With grid bias negative, the current mode with a signal applied to the base is best. The reason for this will become obvious as the parameters are discussed.

BASE-COLLECTOR SHORT-CIRCUIT FORWARD CURRENT TRANSFER RATIO: More commonly known as the common emitter beta, this parameter is the dominant factor in determining available transducer current gain when the grid bias is negative. The variation of beta as a function of bias conditions is shown in figures 8, 9, and 10. Figure 8 shows beta for several values of collector current as a function of grid bias. For a grid bias more than a few volts negative, beta becomes constant with respect to grid bias. The tetrode loses the grid as a useful terminal and takes on the characteristics of a normal transistor. The explanation is found in figure 5.

Assuming that there had been a channel formed as in figure 5d, the effect of a negative grid voltage is to raise the energy levels in the p-region at the surface thus reducing the channel or inversion layer

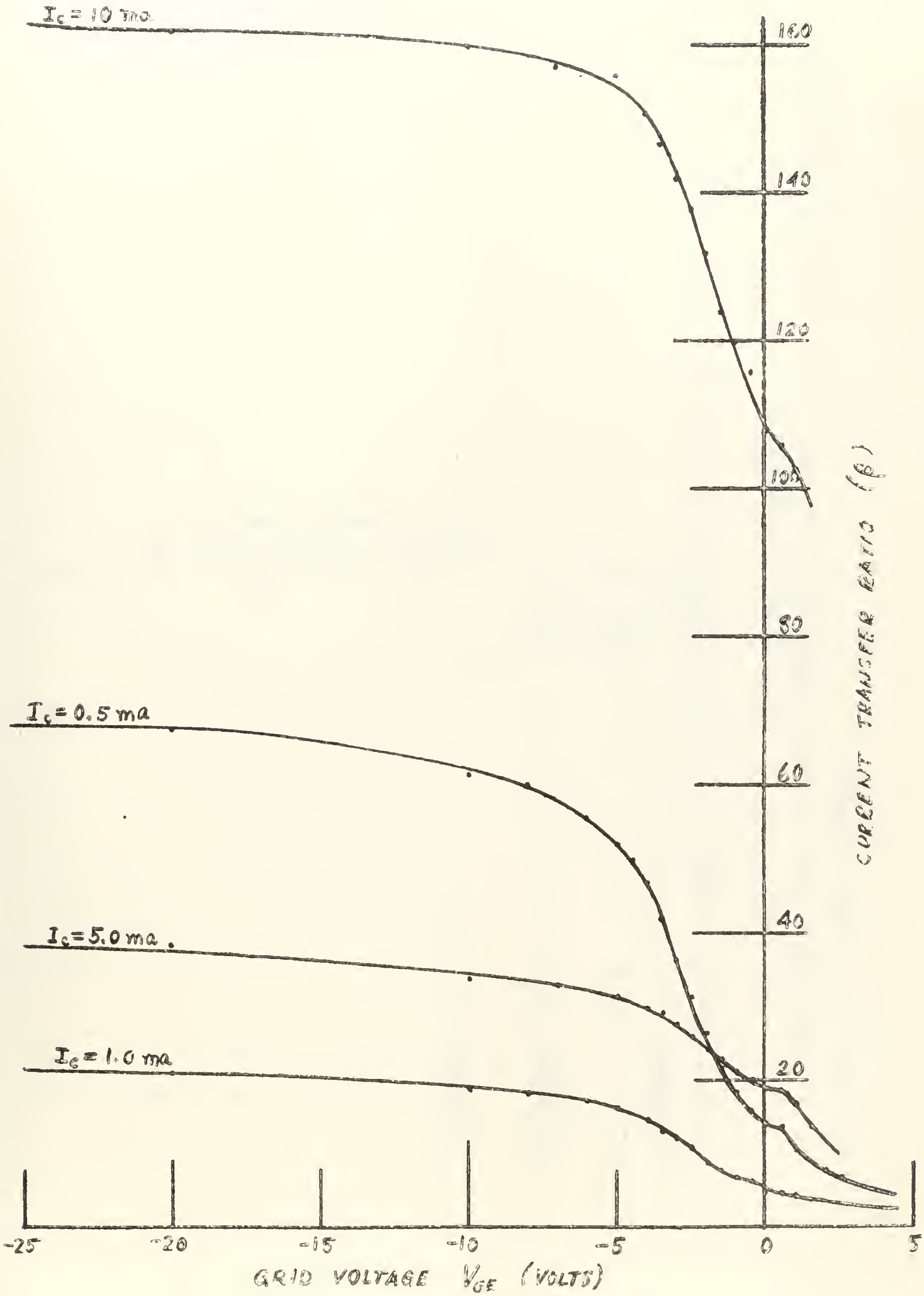


Fig. 8 Common emitter current transfer ratio as a function of V_{ge}

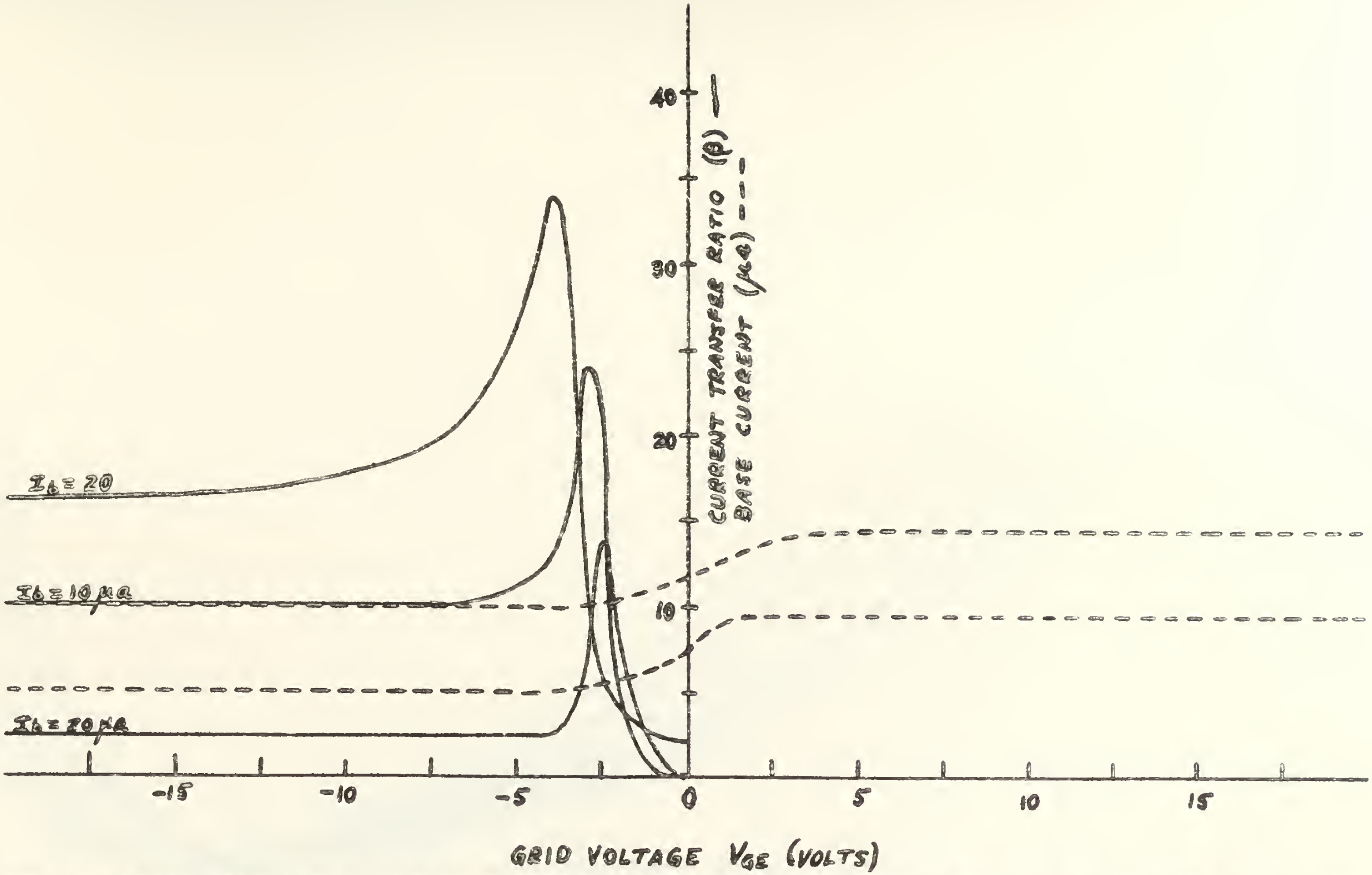


FIG. 10. BASE CURRENT AND COMMON EMITTER CURRENT TRANSFER RATIO AT LOW COLLECTOR CURRENT AS A FUNCTION OF GRID VOLTAGE

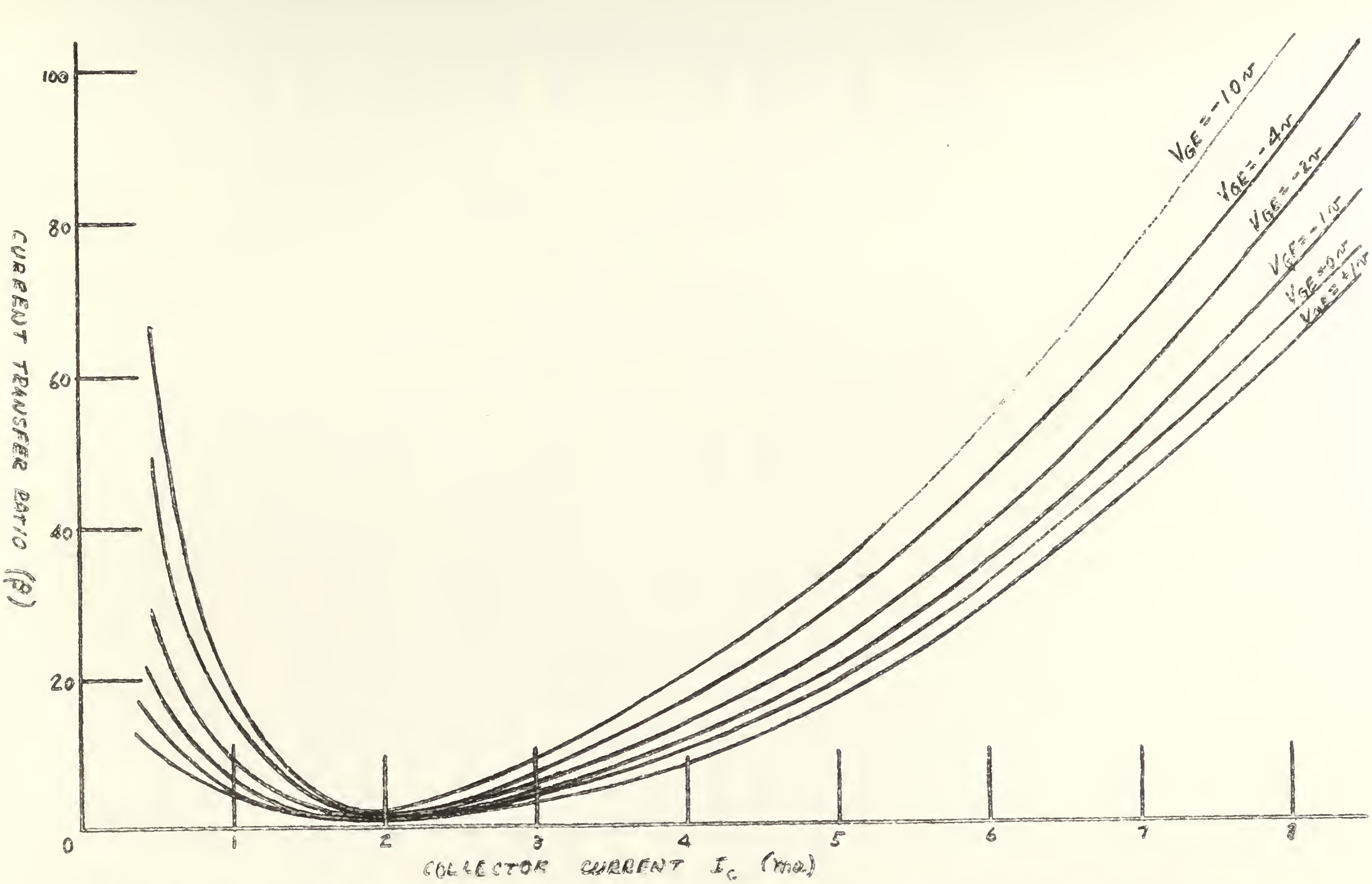


FIG. 9 COMMON EMITTER CURRENT TRANSFER RATIO AS A FUNCTION OF COLLECTOR CURRENT

electron concentration. Applying a still more negative voltage will restore the conditions of figure 5b and eventually bend the levels upward until the junction at the surface is less biased than the bulk. The point of maximum recombination will shift further out of the junction until further increases in grid voltage do not significantly change the rate of recombination in the junction. This accounts for the leveling off of beta with grid voltage well negative.

When the grid voltage is made positive a channel is formed under the grid as shown in figure 5d and a region of high recombination is formed. The recombination in this region reduces the number of minority carriers which arrive at the collector. Since both the rate of recombination in the channel and the carrier concentration in the channel are functions of the emitter-base junction voltage, a time varying junction voltage causes a time varying recombination current which swings about the recombination current determined by the DC grid bias. This is equivalent to a loss of signal in the form of signal component of recombination. Thus, as grid voltage is made positive, both the a-c and the d-c beta are reduced. The terminal value of beta for large positive grid voltage is, of course, zero since it is possible to cut the collector off by means of the grid voltage.

The anomalous bump in the curves of figure 8 in the vicinity of $V_{GE} = 0$ appear to coincide with the onset of channel and the almost immediate saturation of the depth of the channel.

Figure 9 displays the variation of beta as a function of collector current. It is typical of transistors to have an increasing beta vs. collector current characteristic, however it is a montonic function. The curves of figure 9 have an anomalous low current behavior, the cause of which is not clear. In figure 9 the collector current was maintained

at a constant value by varying the base bias. In figure 10 the base was set at a fixed voltage and the grid voltage alone was varied. The collector current and base current were allowed to float. In this instance beta exhibited not only the characteristics of figure 9 but also the peaking shown in figure 10 prior to being cut off by positive grid voltage. These characteristics are obviously associated with the onset of channel and the characteristics of the channel. But it is not explainable in terms of the simple model being used here. The polarity of beta is such as to give a 180° phase shift and so β_{BC} is positive according to the convention of figure 7.

GRID-COLLECTOR TRANSCONDUCTANCE: The characteristic which sets the tetrode apart from the present family of transistors is that, like vacuum tube small signal amplifiers, it draws no power from the signal source. This is brought about by the insulation of the surface of the emitter-base diode by the silicon dioxide grid and the effect on collector current of an electrostatic field imposed on this oxide layer. In terms of equivalent circuits this is a voltage controlled current source in the collector and describes a transfer characteristic or transconductance. Figures 11 and 12 show the variation of the grid transconductance.

Figure 12 plots transconductance as a function of collector current. From zero at collector cutoff, the transconductance increases monotonically with increasing collector current. When collector current increases, it is the result of an increased bias on the emitter-base junction. An increased bias as we have seen changes the character of the channel and the rate of recombination in the channel. Specifically, the position of the point of maximum recombination shifts within the channel. Since carrier concentration, conductivity and recombination rate are all functions the surface potential (and hence base bias) the

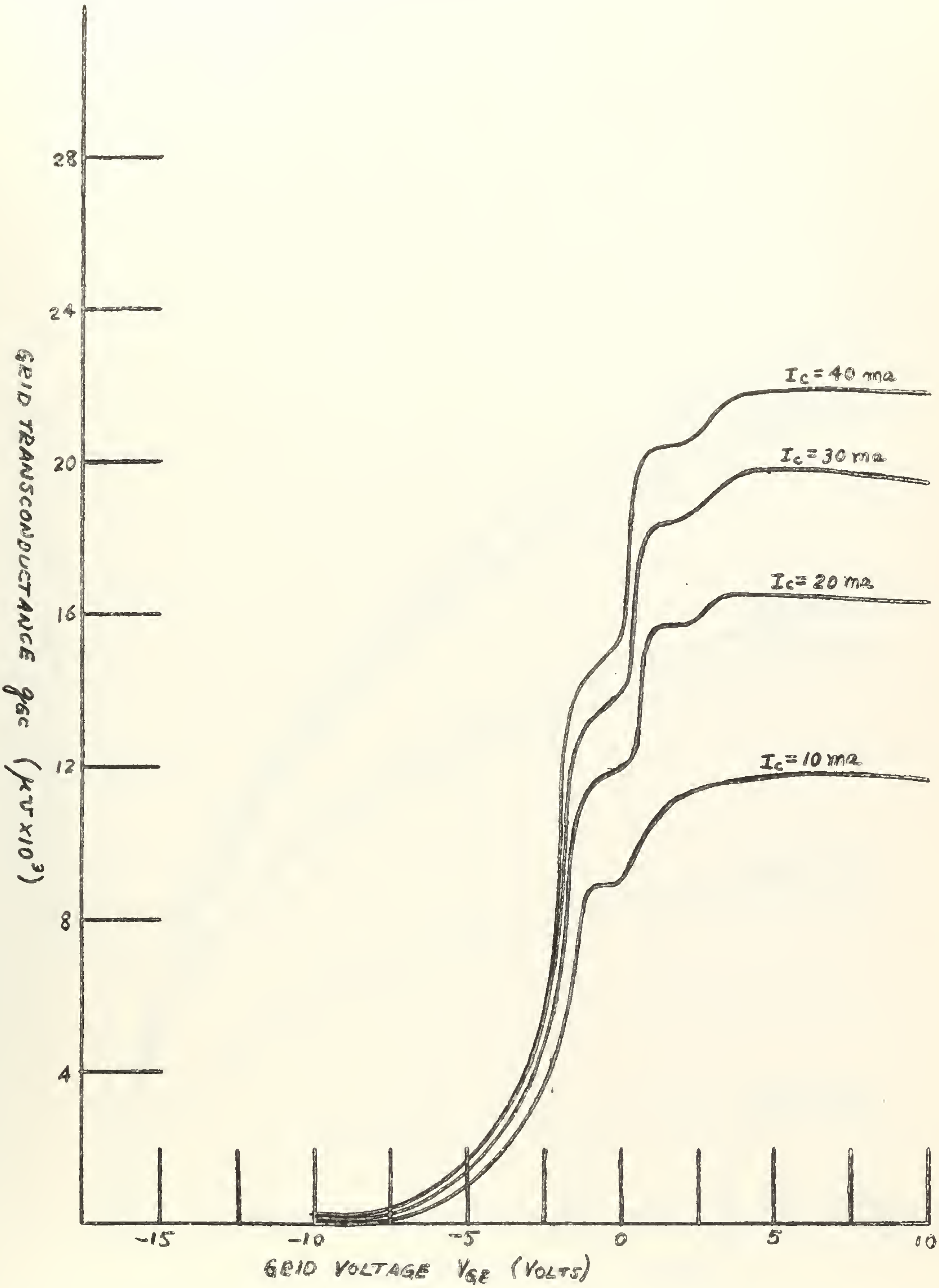


Fig. 11 Grid transconductance as a function of grid voltage

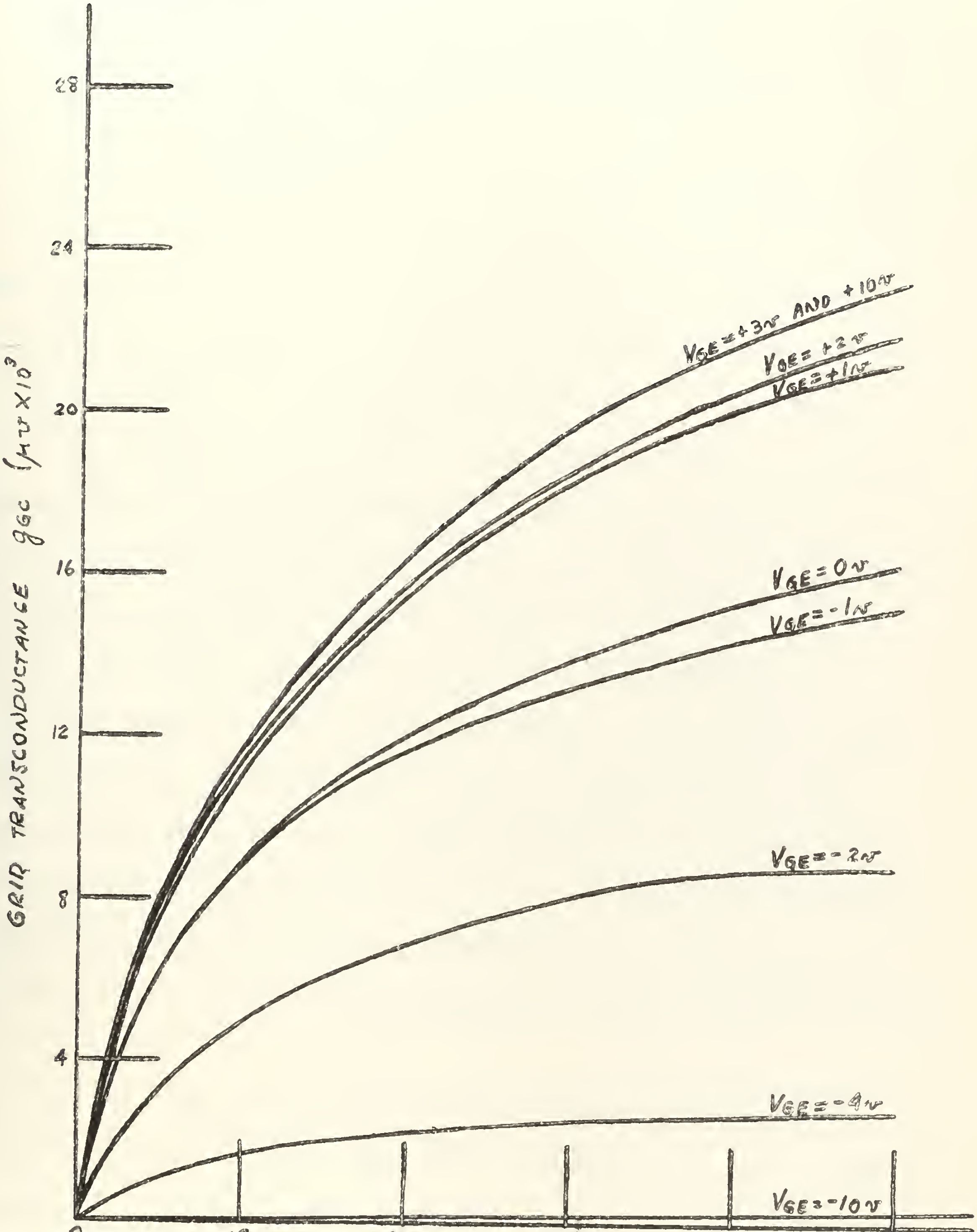


Fig. 12 Grid transconductance as a function of collector current

total recombination current changes, as well as the slope of the recombination current vs. potential characteristic. Since the slope of this characteristic determines the change in collector current for a change in grid voltage, in effect it determines the transconductance of the grid.

The change in transconductance with grid voltage is given in figure 12. Here it is noticed that the antithesis of the beta characteristic takes place. For a negative grid voltage the channel is eliminated entirely and a time varying small signal on the grid will produce no change in collector current. The transconductance is zero. As grid bias is increased from a large negative value toward positive values, a channel forms and the characteristic recombination current begins to flow. If a time varying signal is superimposed upon the grid bias, the recombination current will be modulated with this same time variation. If the base is biased from a d-c source, the modulation of the recombination current will appear as a signal in the collector circuit. The transconductance increases rapidly and then levels off, indicating that the operating point is on the broad, fairly straight side of the recombination rate curve. (It is the slope of the recombination rate curve which predominantly determines the transconductance.) In terms of figure 7, g_{GC} is negative since there is no signal phase shift from grid to collector.

As in the case of beta, an anomalous behavior around zero volts is observed. However there are two anomalies instead of the one previously encountered. There are several plausible explanations which involve the levels of recombination centers (traps), recombination rate vs. surface potential and the onset of channel and its saturation in depth. The real explanation is probably some admixture of all of these but in any

event, any such conclusion would not be supported by experimental evidence in this paper.

The transconductance appears to reach a maximum at approximately five volts and then starts a slow decline. An inflection point in the recombination current characteristic has been reached and the slope decreases as the recombination current tends toward a maximum.

It is interesting to compare the magnitude of transconductance of the tetrode with some typical vacuum tubes:

12DS7	15,000 μmhos
12DA8	15,000 μmhos
12DL8	15,000 μmhos
TETRODE . . .	20,000 μmhos

The semiconductor tetrode compares favorably as can be seen. Its obvious limitation is its power handling capabilities although the free air dissipation of 0.8 watts and a collector breakdown voltage of the order of 100 volts speaks well for the tetrode.

GRID ADMITTANCE, yc: Coupled with the ability to control collector current by means of grid voltage, the impedance level of the grid is a notable characteristic putting the tetrode in a category of electron devices which includes vacuum tubes. As announced by Dr. Sah, the input admittance was that of a very low loss capacitor of the order of 80 pf. There was no change in admittance levels noticeable with the change of any d-c bias condition. No change in the loss coefficient was noticeable from d-c to 10 megacycles. This property is, of course, due to the very good insulating properties of the silicon dioxide layer which grows during the manufacturing process. Such a high input impedance opens areas of circuit applications here-to-fore held to be the sacred domain of vacuum tubes. Included are such applications as operational amplifiers, PAM

demodulators and electrometers.

BASE IMPEDANCE, Z_B: In the common-emitter configuration, the base input impedance is usually thought of as being made up of two resistors in series. The first is r_{bb}', the ohmic resistance of the base bulk material and the second is the resistance associated with the active junction. In the Tee equivalent of the common emitter configuration the resistance of the active junction is composed of that portion through which only the base current flows and that through which the emitter current flows, that is, r_b' and r_E respectively. Since it is always true that $I_E = (\beta+1) I_B$, the voltage drop $I_E r_E$ which is common to both the base and the emitter circuits may be represented in the base circuit as $(\beta+1) I_B r_E$, an equivalent resistance $(\beta+1)$ times as large. Now r_b' is the resistance of the forward biased emitter-base diode and may be expressed as:

$$r_b' \doteq \frac{26}{I_E(ma)} \text{ ohms}$$

Neglecting the reactive component of the base input impedance caused by diffusion capacitance, the expression for the resistance can be written as:

$$Z_B = r_{bb}' + \frac{26}{I_E(ma)} + (\beta+1) r_E \qquad (1)$$

This expression contains within itself the explanation for the observed variation of Z_B with bias conditions shown in figures 13 and 14. In figure 13 it is observed that for low current the base impedance is several thousand ohms but drops very rapidly to a saturation level for currents above a milliamp. Using collector current as an approximation to emitter current, it is seen that the second term of (1) is very large for very low current but drops to 26 ohms at one milliamp. This is the factor which causes rapid changes in resistance with collector current. Notice that the curves for negative grid voltage do not drop as rapidly

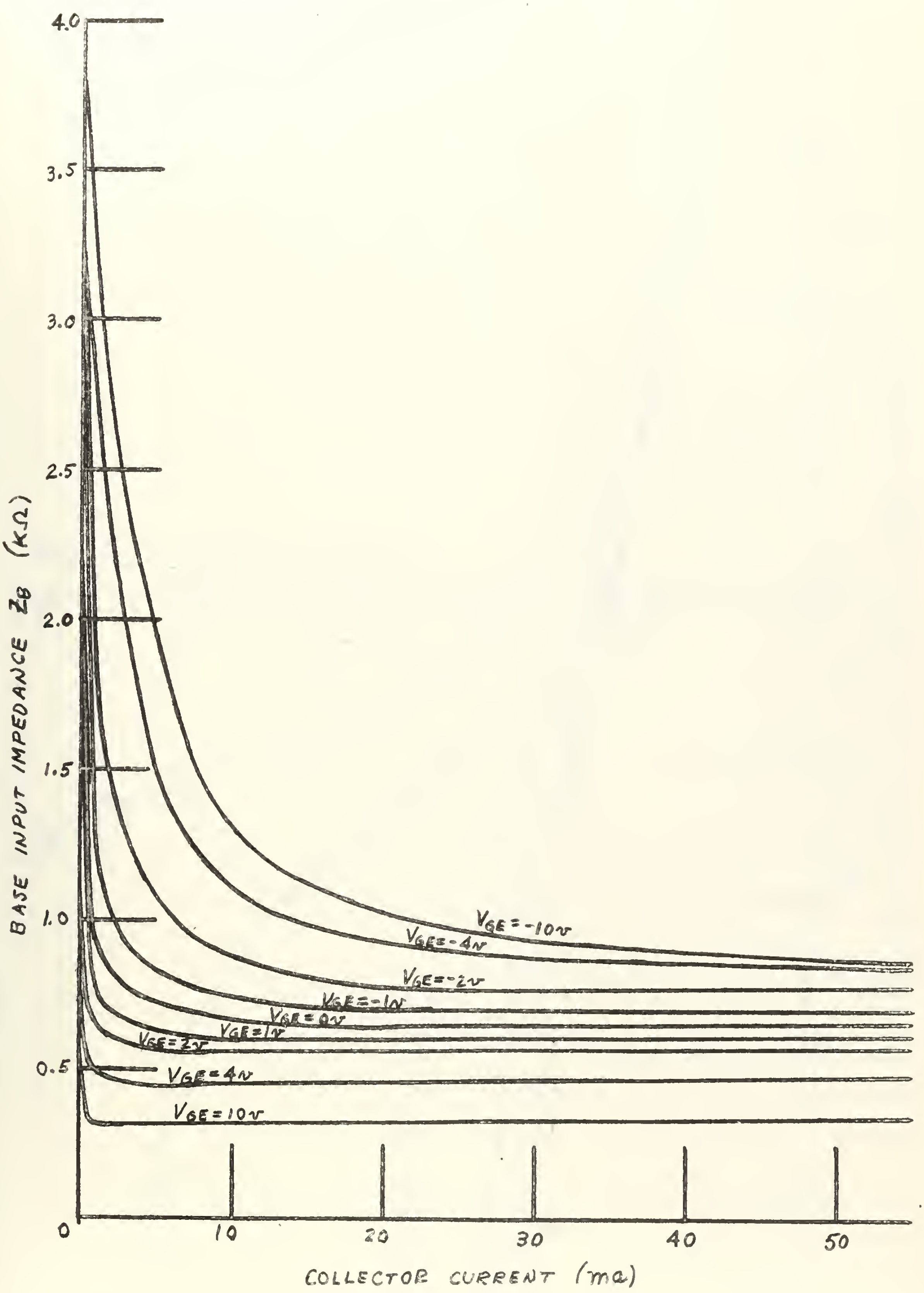


Fig. 1? Base resistance as a function of collector current

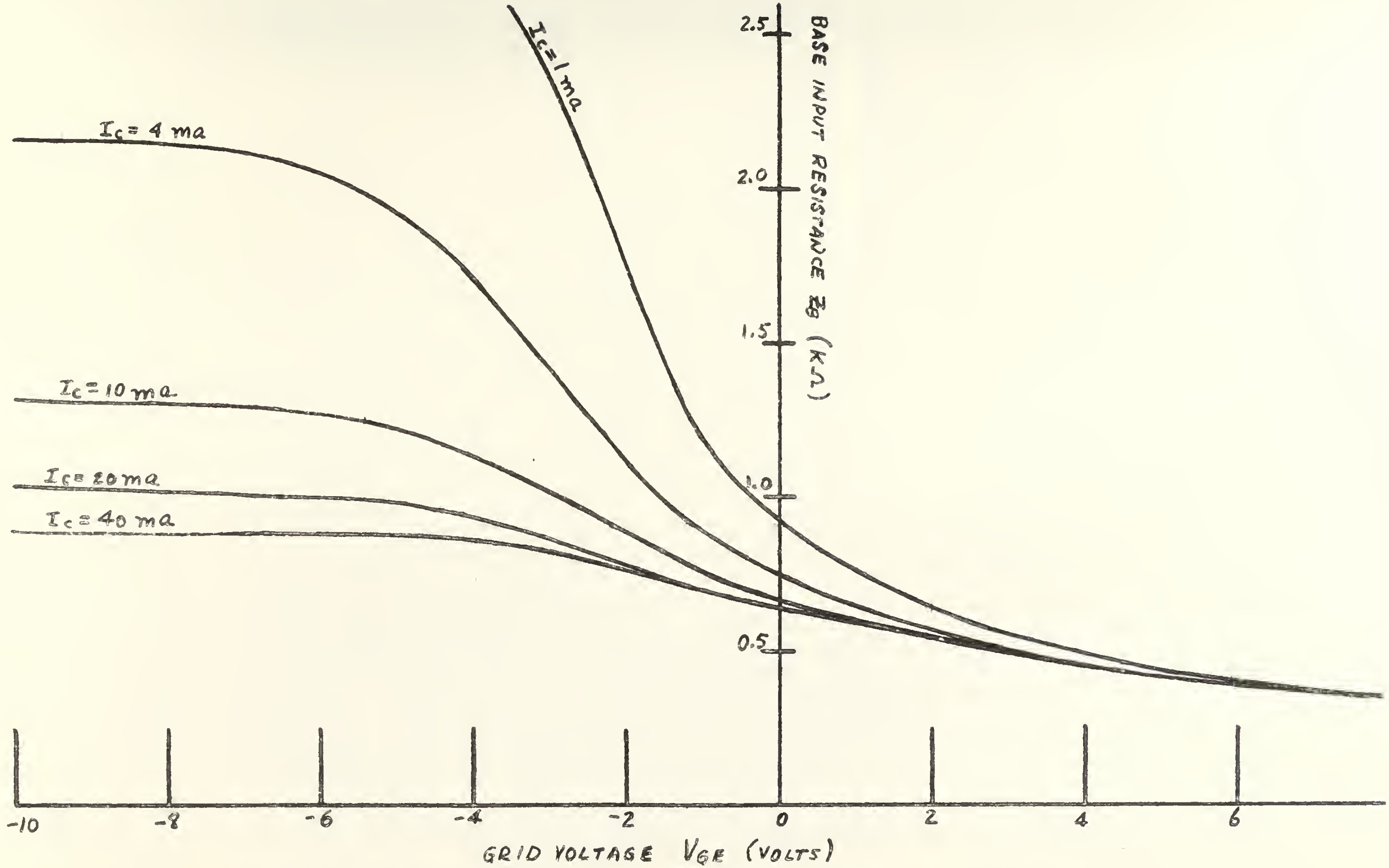


FIG. 14 BASE INPUT RESISTANCE AS A FUNCTION OF GRID VOLTAGE

ror dc [illegible] drop is low. This is the effect of the third term of (1). It may be recalled that β is very high for negative grid voltages but is virtually zero for positive grid voltages. The third term is significant for negative grid voltages and for this case it is dominant when collector current is above one milliamp. As the grid voltage swings positive, a channel is formed and β decreases drastically, approaching zero as grid bias is increased. When this happens, both the second and third terms are negligible and only $r_{bb'}$ remains. That this actually occurs is suggested by the merging of the curves when grid voltage is positive in figure 14.

COLLECTOR-BASE FEEDBACK RATIO: The collector-base feedback ratio in the tetrode is analogous to the same parameter in ordinary transistors. In the common emitter configuration this internal feedback is positive. The cause of the feedback is the modulation of the base width by collector signal voltage. A positive change in collector voltage due to a signal causes the collector junction to spread into the base. This causes an increase in concentration gradient and hence an increase in emitter current. Since the emitter current is directly related to the base voltage by the diode equation, this change causes an increase in forward bias of the emitter-base junction. Thus it can be seen to be a positive voltage feedback in the common emitter configuration. This is also true of the tetrode. Figures 15 and 16 show the variation of collector-base feedback ratio as functions of collector current and grid current respectively.

For low collector currents and negative grid voltages, the feedback is relatively large, decreasing rapidly to a stable value at medium and high currents. This is characteristic behavior for this parameter in

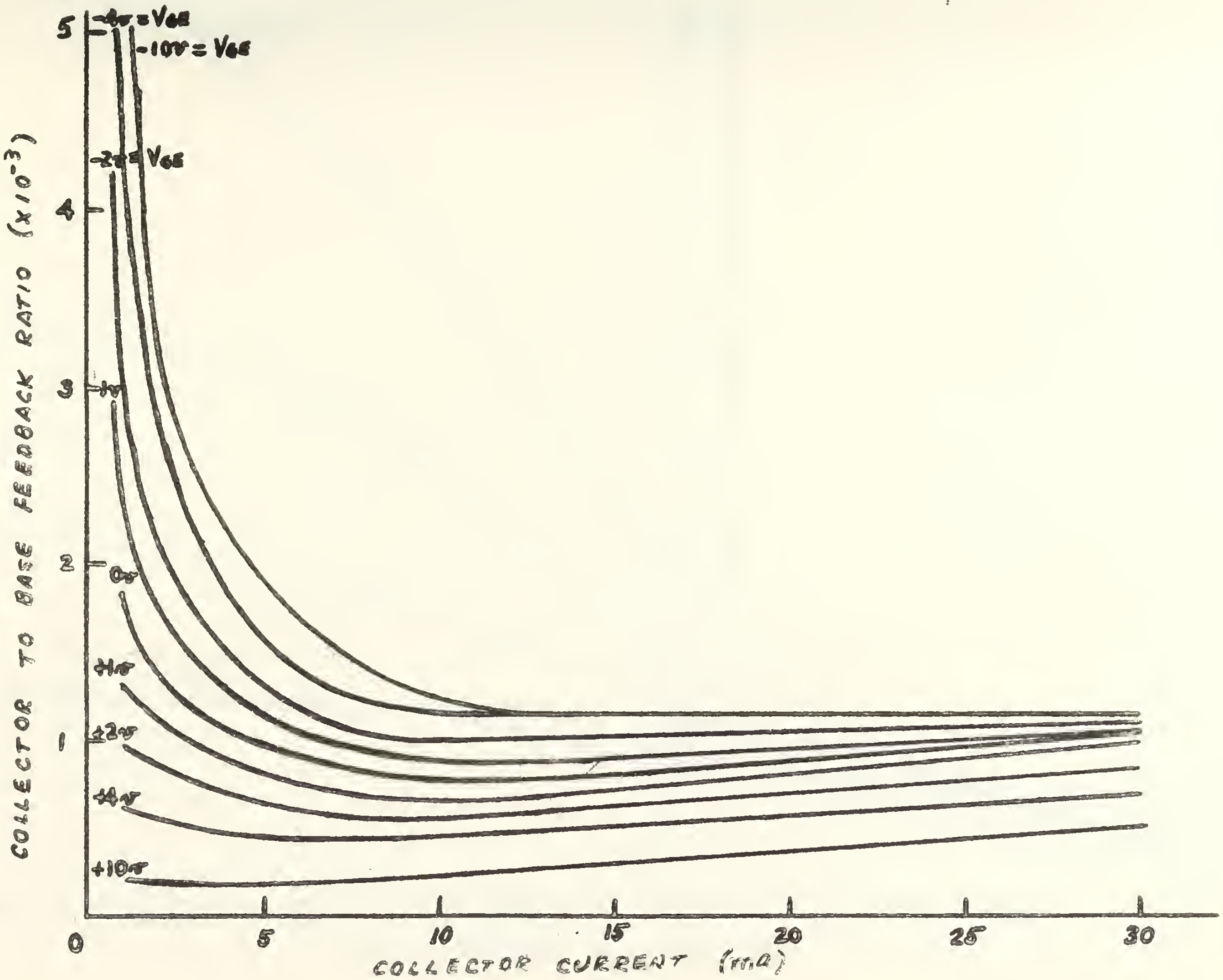


FIG. 15 COLLECTOR TO BASE FEEDBACK RATIO AS A FUNCTION OF COLLECTOR CURRENT

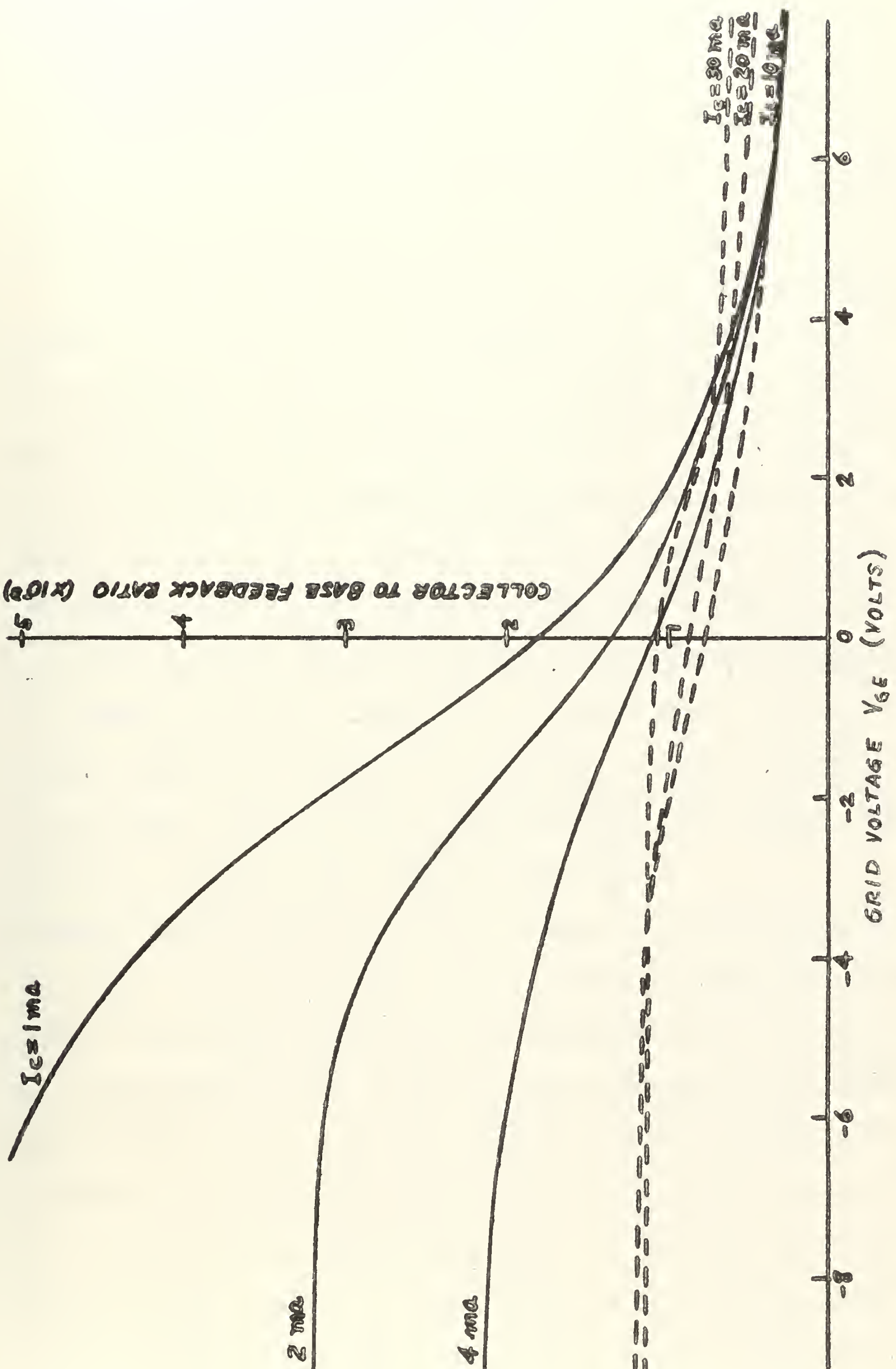


FIG. 16 COLLECTOR TO BASE FEEDBACK RATIO AS A FUNCTION OF GRID VOLTAGE

ordinary transistors* which reinforces the notion that the tetrode behaves in the usual transistor manner when grid voltage is negative. Once collector current gets above about 10 ma, the feedback ratio is insensitive to either grid voltage or further increases in collector current.

When the grid voltage becomes positive and the surface channel is formed, the feedback ratio drops regardless of collector current. It can be seen from figure 16 that for grid voltage greater than two volts, the feedback parameter is essentially constant. In this respect, the surface channel acts in a manner analogous to a surge tank in a fluid system. As the change of collector voltage changes the concentration gradient in the base, this change is reflected in a change in the recombination rate in the channel preferentially to a change in bias across the emitter-base diode in the bulk. In essence then, the change in collector voltage may be thought of as causing two opposing feedback effects: (1) the usual internal feedback (positive) and; (2) a negative feedback which is a function of grid voltage (negative).

GRID-BASE FEEDBACK RATIO: This ratio measures the effect of a signal on the grid in the base circuit. With grid voltage negative and no channel formed, the feedback is that which is associated with the capacity of the metal-oxide-semiconductor sandwich. With increasing grid voltage the surface channel forms. A signal applied to the grid modulates the recombination current in the channel. Since this current reduces the current crossing the emitter-base junction, it is reflected as a reduced junction voltage. Since a positive change in grid voltage decreases junction voltage, the feedback is negative. In part, this feedback may

*See, for instance, GE Transistor Manual, Fourth Edition

[illegible] negative feedback component of [illegible] collector voltage is [illegible] the grid is used as a signal terminal.

A general trend is apparent from the discussion above. Large negative grid voltages eliminate surface channel and makes the tetrode a conventional transistor. The grid has no appreciable control of characteristics. As grid voltage is increased, a surface channel is formed under the grid on the base side of the emitter-base junction and a region of high charge carrier recombination forms. This has the effect of shorting out a signal on the base whether it be intentionally applied or fed back internally. However this same recombination current changes the collector current. Since the recombination current is increased by increasing grid voltage, collector current is decreased. A varying grid voltage now becomes the vehicle for varying collector current in the tetrode so the grid is now the control terminal instead of the base. Since the grid voltage controls collector current, it is now a voltage controlled device, whereas the base was used in the usual current drive mode. The variation of grid or base feedback ratio is shown as a function of grid voltage in Figure [illegible].

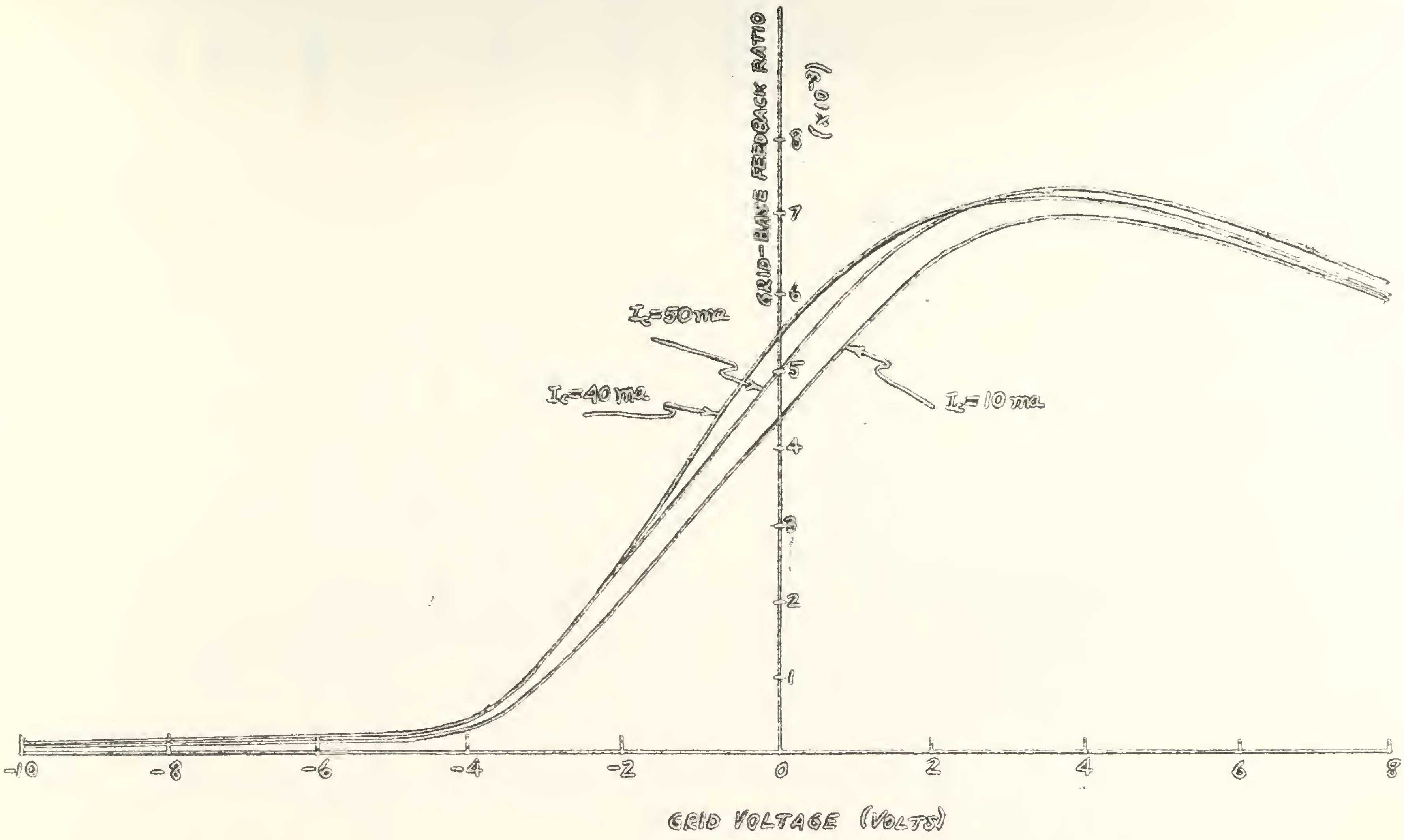


FIG. 17 GRID TO BASE FEEDBACK RATIO AS A FUNCTION OF GRID VOLTAGE

8. Gain and Impedance Equations

The most important characteristics describing an active device from the circuit designer's standpoint are transducer gain and the impedance levels of the input and output terminals. In the common emitter configuration which is being considered here, there are two logical input terminals: the grid and the base; and one logical output terminal: the collector. The base is typically a low impedance terminal and so is associated with a current driven mode of operation. The grid with its very high input impedance and its analogy to vacuum tubes is appropriately the input terminal to be used when driving from a low impedance source. Of importance then is the impedance of the base, grid and collector terminals and the base to collector and grid to collector gain equations. In the former case we are interested in a current gain while in the latter, a voltage gain.

Referring to figure 18 the current-voltage relationship of the base may be written as:

$$V_1 = I_1 Z_B + \mu_{CB} V_2 + \mu_{GB} V_3$$

Since the base input impedance is the ratio of V_1 to I_1,

$$Z_{IN} = \frac{V_1}{I_1} = Z_B + \mu_{CB} \frac{V_2}{I_1} + \mu_{GB} \frac{V_3}{I_1}$$

It is quite apparent that the impedance of the base terminal is not independent of the conditions which exist at the grid and collector. Assume that there is a source of bias and a signal power in the base and grid, each with an associated impedance, and a load impedance in the collector current as shown in figure 18. From this equivalent circuit it is possible to derive an exact expression for the input impedance to the base terminals in the common emitter configuration. The derivation of this

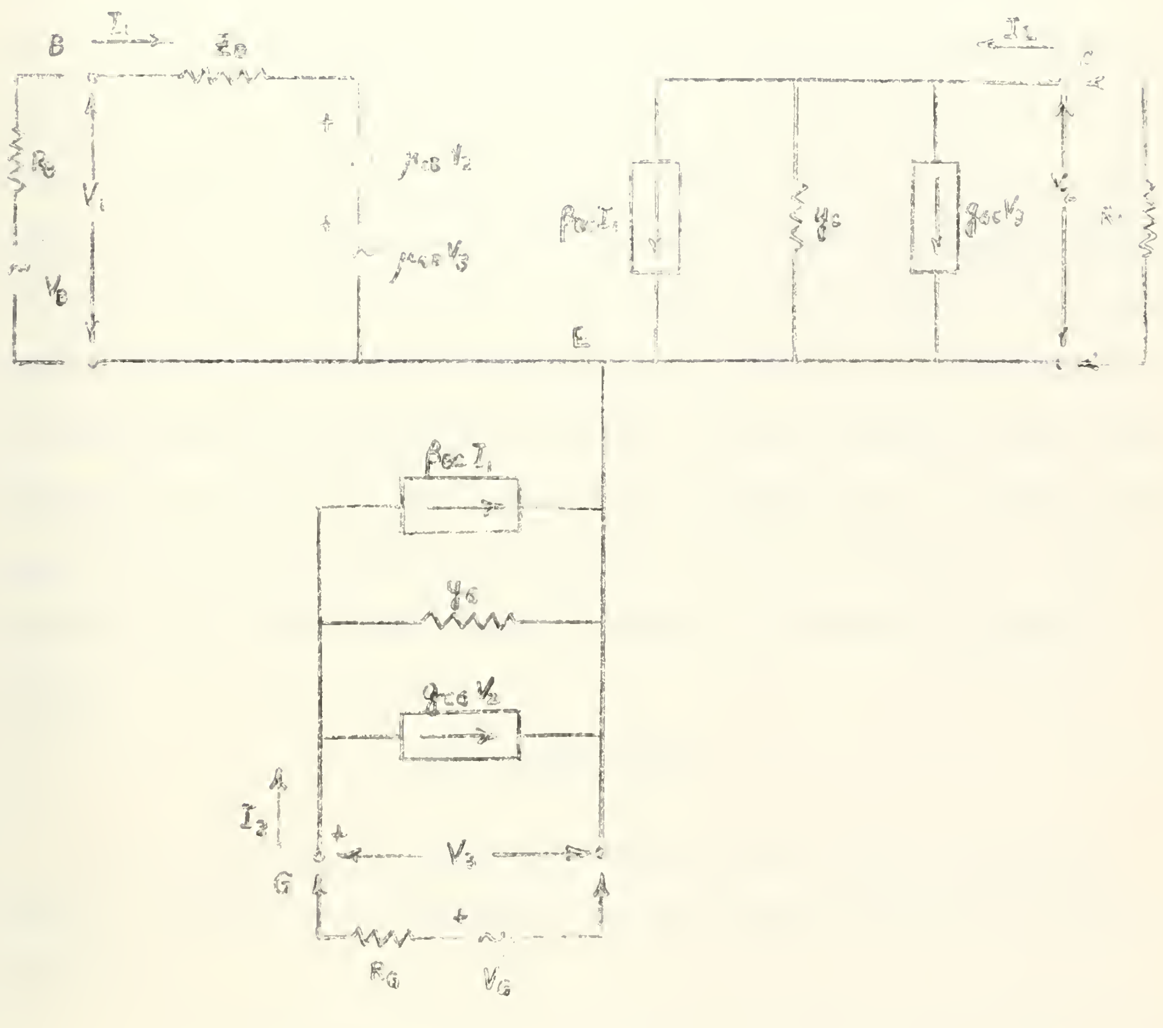


Fig. 18 Tetrode common emitter small signal linear equivalent circuit

expression is derived in Appendix 1

$$Z_{IN} = Z_B + \left[\frac{\mu_{CB} R_L g_{GC}}{R_L y_C + 1}\right] + \left[\frac{\mu_{CB}(R_L y_C + 1) + \mu_{CB} \beta_{BC} R_L}{R_L y_C + 1}\right]\left[\frac{\frac{V_G}{I_G} + \left(\frac{R_L R_G g_{GC}}{R_L y_C + 1}\right) + \beta_{BC} R_G}{1 + R_G y_G - R_G R_L g_{CG} g_{GC}}\right]$$

For negative grid bias, it was seen that Z_B was of the order of several thousand ohms. The second term should be of the order of $R_L/5$ since β_{BC} is approximately 200 and μ_{CB} of the order of 10^{-3} for reasonable collector currents. (Common emitter output impedances, $1/y_C$, have been reported to be in the vicinity of 70K Ω. In most applications $1 + y_C R_L$ will be of the order of unity.) Since a signal would be applied to the base and the grid maintained at a-c ground, the V_G/I_G and $\beta_{BC} R_G$ terms in the numerator of the last part of the Z_{IN} expression would be very small. g_{GC} is approximately zero for negative grid voltage so the denominator of the last bracketed expression is unity. The product $y_G R_G$ is nearly zero since y_G is essentially zero and R_G is the impedance of a d-c bias source. μ_{GB} is essentially zero for this case as is shown by figure 17. Hence for the case of large negative grid voltage, Z_{IN} reduces to approximately:

$$Z_{IN} = Z_B + R_L/5$$

In the voltage mode with large positive grid voltage, the base would be operated at a-c ground and the base input impedance is of no importance.

The grid input admittance is derived in Appendix 2 and is given by:

$$Y_{IN} = y_G + \left[\frac{g_{GC} g_{CG} R_L}{R_L y_C + 1}\right] + \beta_{BC}(R_L y_C + R_L g_{CG} + 1)\left[\frac{\frac{V_B}{V_G} - \mu_{GB} - \left(\frac{\mu_{CB} R_L g_{GC}}{R_L y_C + 1}\right)}{(R_L y_C + 1)(R_B + Z_B) + \mu_{CB} R_L \beta_{BC}}\right]$$

This expression is useful when the tetrode is being used in the voltage, or grid input, mode. For this condition the grid bias is large and

[illegible] β_{BC} [illegible] only [illegible] information on β_{BG} [illegible] form:

$$Y_{IN} \doteq y_G + 20\lambda\, g_{CG}$$

The output admittance expression [illegible] derived in Appendix 3 [illegible] given by:

$$Y_{OUT} = y_C + \left[\frac{\beta_{BC}\left(\frac{V_B}{V_2} - \mu_{CB}\right)}{R_B + Z_B}\right] + \left[g_{BC}\left(R_B + Z_B\right) - \beta_{BC}\mu_{GB}\right]\left[\frac{\frac{V_G}{V_2} + R_G g_{CG} + R_G \beta_{BC}\left(\frac{\frac{V_B}{V_2} - \mu_{CB}}{R_B + Z_B}\right)}{\left(R_G y_G + 1\right)\left(R_B + Z_B\right) + \mu_{GB} R_G \beta_{BG}}\right]$$

Although this expression may be simplified for the two modes of operation, the essential features are that the output admittance is essentially controlled by y_C. There is a term which is dependent both on the magnitude of gain and signal size in both modes of operation. In the current mode V_G and R_G are essentially zero in the third term. $\beta_{BC} \cdot \frac{V_B}{V_2}$ is essentially unity; $\beta_{BC} \cdot \mu_{CB}$ is approximately 0.2 for I_C = 10ma and $R_B + Z_B$ is of the order of 10^4. With these approximations, Y out becomes:

$$Y_{OUT} \doteq y_C + 80\,\mu\mho$$

In the voltage mode β_{BC} and V_B are approximately zero. Since no information is known about β_{BG} no further simplification can be made beyond:

$$Y_{OUT} \doteq y_C + g_{BC}\left[\frac{\frac{V_G}{V_2} + R_G g_{CG} + R_G \beta_{BC}\left(\frac{-\mu_{CB}}{R_B + Z_B}\right)}{\left(R_G y_G + 1\right)\left(R_B + Z_B\right) + \mu_{GB} R_G \beta_{BC}}\right]\left(R_B + Z_B\right)$$

As before, y_C is seen to set the gross value of the output admittance with a gain dependent term modifying this value.

The voltage mode gain expression is given by:

$$A_v = \left[\frac{\beta_{BC}\mu_{GB} R_L - g_{GC}(R_B + Z_B) R_L}{(R_B + Z_B)(R_L y_C + 1) - \beta_{BC}\mu_{CB} R_L}\right] - \left[\frac{\beta_{BC} R_L}{(R_B + Z_B)(R_L y_C + 1) - \beta_{BC}\mu_{CB} R_L}\right] \frac{V_c}{V_g}$$

and is derived in Appendix 4. Large positive grid voltages drive to almost zero which immediately simplifies this expression to:

$$A_v \doteq - \frac{g_{GC} R_L}{R_L y_C + 1}$$

This is identical to the vacuum tube expression for gain. It must be remembered that there is no phase shift through the tetrode so g_{GC} is a negative number according to the completely arbitrary polarities assigned in figure (18).

Current gain in the current operated mode is given by:

$$A_i = \left[\frac{\{\beta_{BC}(R_G y_G + 1) + g_{GC}\beta_{BG} R_G\}(R_L y_C + 1)}{(R_G y_G + 1)(R_L y_C + 1) - g_{CG} g_{GC} R_L R_G}\right] - \left[\frac{g_{GC}(R_G y_G + 1)}{(R_G y_G + 1)(R_L y_C + 1) - g_{GC} g_{CG} R_G R_L}\right] \frac{V}{I}$$

Figure 11 shows that g_{GC} goes to zero for large negative grid voltages. This defines the current mode and current gain simplifies to:

$$A_i \doteq \beta_{BC}$$

The expressions given above describe the tetrode in its two usual modes of operation for low frequency, small signal operating conditions. They should provide the circuit designer with an insight into how the tetrode will respond for a given circuit application.

9. Miscellaneous Notes and Observations

The tetrode exhibited two characteristics which, if not kept in mind, might destroy the device or prevent a circuit from operating properly. The first of these is the effect of base current on total power dissipation and the second is an observed bias point instability.

The tetrode collector current may be cut off by the application of a sufficiently high positive grid voltage. This does not mean, however, that base and emitter current is not flowing. Indeed, quite a large recombination current may be flowing in the channel. This can be several hundred milliamperes under some conditions. The potential difference through which this current flows is the emitter base voltage. The product of the two is the power being dissipated as a result of recombination and this appears as lattice vibration or kinetic energy of the crystal atoms. If this power exceeds the dissipation capability of the device, destruction will result; and without ever having drawn any power at all from the collector source. Thus it becomes apparent that care must be taken to monitor collector current, base current, emitter-base voltage and collector-base voltage. The total power being dissipated is then:

$$P = V_{BE}(I_B+I_C) + V_{CB}I_C$$

with I_B no longer negligible. Contrast this with the old rule of thumb: $P = V_{CE}I_C$. It is obvious that the circuit designer will have to be aware of this pitfall and mindful of the actual physics of this device rather than relying upon the "rule of thumb" methods of design which seem to have grown up with the transistor.

Bias point instability was first observed when the low collector current a-c beta was being investigated. It was noticed that for a given

collector supply voltage, collector current and collector impedance, changes in the grid voltage would change the collector current. The initial conditions could be restored by changing the base current. While not an unexpected result on the surface, it presented an undesirable feedback factor unless it could be eliminated. From an a-c standpoint it was found that the feedback could be eliminated by the addition of a by-passed resistor in the emitter lead. This is the same approach as is used in self bias networks in both tubes and transistors. However, the philosophy of why it works is not the same.

Essentially, changing the grid voltage changes the d-c beta and hence I_C. But changing I_C results in a change in the common emitter input resistance as was explained in section 7

$$Z_B = r_{bb}' + \frac{26}{I_E(ma)} + (\beta+1) r_E = r_B + (\beta+1) r_E$$

If a resistance, R_E, is placed in series with the emitter, this expression becomes:

$$Z_B = r_B + (\beta+1)(r_E + R_E)$$

Since $\beta I_B = I_C$, it follows that for initial conditions $\beta_1 I_{B1} = I_{C1}$ and for final conditions (after grid voltage changes), $\beta_2 I_{B2} = I_{C2}$. If the grid voltage change increases β, and I_B is from a high impedance source, I_B will not change significantly because Z_B is a small fraction of the total impedance in the base loop. However, if I_B is derived from a low impedance source, I_B can be made to "track" β. This is shown to be the case in the following analysis. The desired result is $I_{C1} = I_{C2}$. Refer to figure 19.

$$R_{B1} = R_S + r_B + (\beta_1+1)(r_E + R_E) = R_S + r_B + r_E + R_E + \beta_1 (r_E + R_E)$$

$$R_{B2} = R_S + r_B + (\beta_2+1)(r_E + R_E) = R_S + r_B + r_E + R_E + \beta_2 (r_E + R_E)$$

$$I_{B1} = \frac{V_B}{R_{B1}} \qquad I_{B2} = \frac{V_B}{R_{B2}}$$

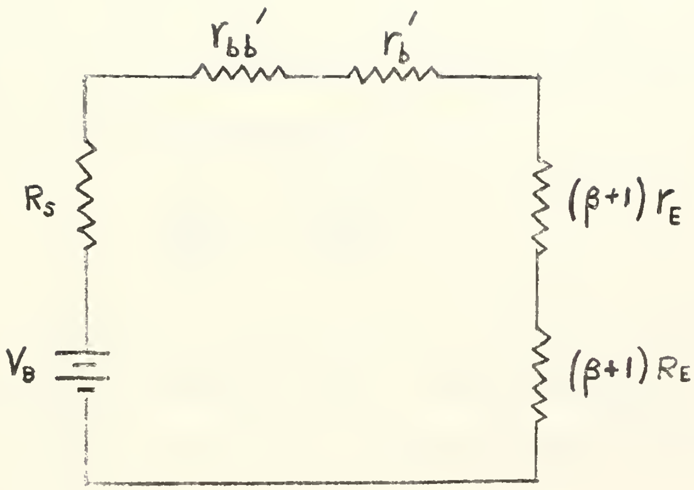


Fig. 19 Common emitter Tee equivalent base circuit neglecting feedback terms

Now let $\beta_2 = \beta_1 + \Delta\beta$ where $\Delta\beta$ is the change caused by a change in grid voltage. Then:

$$I_{B2} = \frac{V_B}{R_{B2}} = \frac{V_B}{R_S + r_B + r_E + R_E + (\beta_1 + \Delta\beta)(r_E + R_E)}$$

and $\beta_1 I_{B1} = \beta_1 I_{B2} + \Delta\beta I_{B2}$ if the desired $I_{C1} = I_{C2}$ is to be achieved. In order to do this, it is necessary to make the ratio I_{B1}/I_{B2} as nearly equal to β_1/β_2 as possible. Therefore it is necessary that:

$$\frac{I_{B1}}{I_{B2}} = \frac{R_{B2}}{R_{B1}} = \frac{R_S + r_B + r_E + R_E + (\beta_1 + \Delta\beta)(r_E + R_E)}{R_S + r_B + r_E + R_E + \beta_1 (r_E + R_E)}$$

be as nearly equal to $(\beta_1 + \Delta\beta)/\beta_1$ as possible. If the d-c beta is reasonably large (~10) and R_s is very small, it can be seen that making R_E large will reduce the above expression to:

$$\frac{I_{B1}}{I_{B2}} \doteq \frac{\beta_1 + \Delta\beta}{\beta_1}$$

as desired. When tested, this technique was observed to almost eliminate the bias instability previously noted. By-passing R_E with a capacitor allowed an a-c signal on the grid to have the desired affect on beta without changing the d-c bias conditions.

10. Summary

A new semiconductor device known as the surface-potential controlled transistor or semiconductor tetrode has been described. The underlying physics have been laid out to aid in the understanding of how it obtains transistor action. The device is characterized by a high impedance terminal to which a potential may be applied to effect control of collector current.

The description of the device in terms of a hybrid set of parameters was given and the behavior of these parameters under varying bias conditions was investigated. Based upon these investigations, the description of the tetrode as a transducer was derived and discussed.

The tetrode is not a commercially available device however, when it is marketed it offers great promise of applications previously barred to semiconductor devices.

BIBLIOGRAPHY

1. C. T. Sah, A New Semiconductor Tetrode--The Surface-Potential Controlled Transistor, Proc. IRE, vol. 49, no. 11, pp. 1623-34, Nov., 1961.

2. A. L. McWhorter and R. H. Kingston, Channels and Excess Reverse Current in Grown Germanium P-N Junction Diodes, Proc. IRE, vol. 42, no. 9, pp. 1376-80, Sept., 1954.

3. W. L. Brown, N-Type Surface Conductivity on P-Type Germanium, Phys. Rev., vol. 91, p. 518, 1 Aug., 1953.

4. J. T. Law, A Mechanism for Water Induced Excess Reverse Dark Current on Grown Germanium N-P Junctions, Proc. IRE, vol. 42, no. 6, pp. 907-13, June, 1954.

5. H. Christensen, Surface Conduction Channel Phenomena in Germanium, Proc. IRE, vol. 42, no. 9, pp. 1371-75, Sept., 1954.

6. J. Bardeen, Surface States and Rectification at a Metal Semiconductor Contact, Phys, Rev., vol. 71, pp. 717-27, May, 1947.

7. M. Cutler and H. M. Bath, Surface Leakage Current in Silicon Fused Junction Diodes, Proc. IRE, vol. 45, no. 6, pp. 39-43, June, 1957.

8. R. H. Kingston and S. F. Neustadter, Calculation of the Space Charge, Electric Field, and Free Carrier Concentration at the Surface of a Semiconductor, J. Appl. Phys., vol. 26, no. 6, pp. 718-20, June, 1955.

9. W. Shockley, Electrons and Holes in Semiconductors, D. Van Nostrand Co., 1950.

10. W. Shockley and W. T. Read, Statistics of Recombination of Holes and Electrons, Phys. Rev., vol. 87, pp. 835-42, Sept., 1952.

11. C. T. Sah, R. N. Noyce, and W. Shockley, Carrier Generation and Recombination in P-N Junctions and P-N Junction Characteristics, Proc. IRE, vol. 45, no. 9, pp. 1228-43, Sept., 1957.

APPENDIX I

DERIVATION OF THE BASE INPUT IMPEDANCE IN THE COMMON EMITTER CONFIGURATION

$V_1 = Z_B I_1 + \mu_{CB} V_2 + \mu_{GB} V_3$ Refer to figure 18

$V_2 = I_2 \mathfrak{z}_2 \qquad \mathfrak{z}_2 = R_L/(R_L y_C + 1) \qquad V_3 = I_3 \mathfrak{z}_3 \qquad \mathfrak{z}_3 = R_G/(R_G y_G + 1)$

$I_2 = \beta_{BC} I_1 + g_{GC} V_3 \qquad I_3 = \dfrac{V_G}{R_G} + g_{CG} V_2 + \beta_{BG} I_1$

$$V_2 = \frac{(\beta_{BC} I_1 + g_{GC} V_3)\, R_L}{R_L y_C + 1} \qquad V_3 = \frac{V_G + R_G g_{CG} V_2 + \beta_{BG} R_G I_1}{R_G y_G + 1}$$

Substituting the expression for V_2 into that for V_3:

$$V_3 = \frac{\left[V_G + \dfrac{R_G R_L g_{CG}}{R_L y_C + 1}\left(\beta_{BC} I_1 + g_{GC} V_3\right) + \beta_{BG} R_G I_1\right]}{R_G y_G + 1}$$

$$\left[1 - \frac{R_G R_L g_{GC} g_{CG}}{R_L y_C + 1}\right] V_3 = \left[\frac{V_G + \left(\dfrac{R_G R_L g_{GC} \beta_{BC}}{R_L y_C + 1}\right) I_1 + \beta_{BC} R_G I_1}{R_G y_G + 1}\right]$$

$$V_3 = \left[\frac{V_G + \left(\dfrac{R_G R_L g_{CG} \beta_{BC}}{R_L y_C + 1}\right) I_1 + \beta_{BG} R_G I_1}{1 + R_L y_C - R_G R_L g_{CG} g_{GC}}\right]\left[\frac{R_L y_C + 1}{R_G y_G + 1}\right]$$

Using this result to solve for V_2:

$$V_2 = \frac{R_L \beta_{BC} I_1}{R_L y_C + 1} + \left[\frac{g_{GC} R_L}{R_G y_G + 1}\right]\left[\frac{V_G + \left(\dfrac{R_G R_L g_{CG} \beta_{BC}}{R_L y_C + 1}\right) I_1 + \beta_{BG} R_G I_1}{1 + R_L y_C - R_G R_L g_{CG} g_{GC}}\right]$$

Substituting these expressions for V_2 and V_3 into the original expression for V_1, it is found to be a function of I_1 alone. Since the ratio of V_1 to I_1 is the input impedance of the base, the solution may be found:

$$V_1 = Z_B I_1 + \frac{\mu_{CE} R_L \beta_{BC} I_1}{R_L y_C + 1} + \left[\frac{\mu_{CB}\, g_{GC} R_L}{R_G y_G + 1}\right]\left[\frac{V_G + \left(\frac{R_G R_L g_{CG} \beta_{BC}}{R_L y_C + 1}\right) I_1 + \beta_{BC} R_G I_1}{1 + R_L y_C - R_G R_L g_{CG}\, g_{GC}}\right] +$$

$$\mu_{GB}\left[\frac{R_L y_C + 1}{R_G y_G + 1}\right]\left[\frac{V_G + \left(\frac{R_G R_L g_{CG} \beta_{BC}}{R_L y_C + 1}\right) I_1 + \beta_{BG} R_G I_1}{1 + R_L y_C - R_G R_L g_{CG}\, g_{GC}}\right]$$

$$Z_{IN} = \frac{V_1}{I_1} = Z_B + \frac{\mu_{CB} R_L \beta_{BC}}{R_L y_C + 1} + \left[\frac{\mu_{GB}(R_L y_C + 1) + \mu_{CB}\, g_{GC} R_L}{R_G y_G + 1}\right] \cdot$$

$$\left[\frac{\frac{V_G}{I_1} + \left(\frac{R_L R_G g_{CG} \beta_{BC}}{R_L y_C + 1}\right) + \beta_{BC} R_G}{1 + R_L y_C - R_G R_L g_{CG}\, g_{GC}}\right]$$

APPENDIX II

DERIVATION OF THE GRID INPUT ADMITTANCE IN THE COMMON EMITTER CONFIGURATION

$$I_3 = \beta_{BG} I_1 + y_G V_3 + g_{CG} V_2$$ Refer to figure 18

$$V_2 = I_2 \, g_2 \qquad g_2 = R_L/(R_L y_c + 1) \qquad I_1 = \frac{V_B - \mu_{CB} V_2 - \mu_{GB} V_3}{R_B + Z_B}$$

$$I_2 = \beta_{BC} I_1 + g_{GC} V_3$$

$$V_2 = \frac{(\beta_{BC} I_1 + g_{GC} V_3)}{R_L y_c + 1}$$

Substituting the expression for V_2 into the expression for I_1:

$$I_1 = \frac{V_B}{R_B + Z_B} - \frac{\mu_{CB} R_L \beta_{BC} I_1}{(R_L y_c + 1)(R_B + Z_B)} - \frac{\mu_{GB} V_3}{R_B + Z_B} - \frac{\mu_{CB} R_L g_{GC} V_3}{(R_L y_c + 1)(R_B + Z_B)}$$

$$I_1 = (R_L y_c + 1)\left[\frac{V_B - \mu_{GB} V_3 - \dfrac{\mu_{CB} R_L g_{GC} V_3}{R_L y_c + 1}}{(R_L y_c + 1)(R_B + Z_B) + \mu_{CB} R_L \beta_{BC}}\right]$$

And using this result to solve for V_2

$$V_2 = \beta_{BC} R_L \left[\frac{V_B - \mu_{GB} V_3 - \dfrac{\mu_{CB} R_L g_{GC} V_3}{R_L y_c + 1}}{(R_L y_c + 1)(R_B + Z_B) + \mu_{CB} R_L \beta_{BC}}\right] + \frac{g_{GC} R_L V_3}{R_L y_c + 1}$$

Substituting these expressions for V_2 and V_3 into the original expression for I_3, it is found to be a function of V_3 alone. Since the ratio of I_3 to V_3 is the grid input admittance:

$$I_3 = \beta_{BC} (R_L y_c + 1) \left[\frac{V_B - \mu_{GB} V_3 - \left(\dfrac{\mu_{CB} R_L g_{GC}}{R_L y_c + 1}\right) V_3}{(R_L y_c + 1)(R_B + Z_B) + \mu_{CB} R_L \beta_{BC}}\right] + y_G V_3 + \frac{g_{CG} g_{GC} R_L V_3}{R_L y_c + 1} +$$

$$g_{CG} \beta_{BC} R_L \left[\frac{V_B - \mu_{GB} V_3 - \left(\dfrac{\mu_{CB} R_L g_{GC}}{R_L y_c + 1}\right) V_3}{(R_L y_c + 1)(R_B + Z_B) + \mu_{CB} R_L \beta_{BC}}\right]$$

$$Y_{IN} = \frac{I_3}{V_3} = y_G + \left[\frac{g_{CG}\, g_{GC} R_L}{R_L y_C + 1}\right] + \beta_{BC}\left(R_L y_C + R_L g_{CG} + 1\right) \cdot$$

$$\left[\frac{\frac{V_B}{V_3} - \mu_{GB} - \left(\frac{\mu_{CB} R_L g_{GC}}{R_L y_C + 1}\right)}{\left(R_L y_C + 1\right)\left(R_B + Z_B\right) + \mu_{CB} R_L \beta_{BC}}\right]$$

APPENDIX III

DERIVATION OF THE COLLECTOR OUTPUT ADMITTANCE IN THE COMMON EMITTER CONFIGURATION

$I_2 = y_c V_2 + \beta_{BC} I_1 + g_{GC} V_3$ Refer to figure 18

$V_3 = I_3 z_3 \quad z_3 = R_G/(R_G y_G + 1)$ $\qquad I_1 = \dfrac{V_B - \mu_{CB} V_2 - \mu_{GB} V_3}{R_B + Z_B}$

$$I_3 = \frac{V_G}{R_G} + g_{CG} V_2 + \beta_{BG} I_1$$

$$V_3 = \frac{V_G + R_G g_{CG} V_2 + \beta_{BG} R_G I_1}{R_G y_G + 1}$$

Substituting the expression for I_1 into the expression for V_3

$$V_3 = \frac{V_G + R_G g_{CG} V_2}{R_G y_G + 1} + \frac{\beta_{BG} R_G}{R_G y_G + 1}\left[\frac{V_B - \mu_{CB} V_2 - \mu_{GB} V_3}{R_B + Z_B}\right]$$

$$V_3\left[1 + \frac{\beta_{BG} R_G \mu_{GB}}{(R_G y_G + 1)(R_B + Z_B)}\right] = \frac{V_G + R_G g_{CG} V_2 + \beta_{BG} R_G \left(\frac{V_B - \mu_{CB} V_2}{R_B + Z_B}\right)}{R_B + Z_B}$$

$$V_3 = \left[\frac{V_G + R_G g_{CG} V_2 + R_G \beta_{BG}\left(\frac{V_B - \mu_{CB} V_2}{R_B + Z_B}\right)}{(R_G y_G + 1)(R_B + Z_B) + \mu_{GB} R_G \beta_{BG}}\right](R_B + Z_B)$$

And using this result to solve for I_1:

$$I_1 = \frac{V_B - \mu_{CB} V_2}{R_B + Z_B} - \mu_{GB}\left[\frac{V_G + R_G g_{CG} V_2 + R_G \beta_{BG}\left(\frac{V_B - \mu_{CB} V_2}{R_B + Z_B}\right)}{(R_G y_G + 1)(R_B + Z_B) + \mu_{GB} R_G \beta_{BG}}\right]$$

Substituting these expressions for V_3 and I_1 into the original expression for I_2, it is found to be a function of V_2 alone. Since the ratio of I_2 to V_2 is the collector output admittance:

$$I_2 = y_c V_2 + \frac{\beta_{BC}(V_B - \mu_{CB} V_2)}{R_B + Z_B} - \beta_{BC}\mu_{GB}\left[\frac{V_G + R_G g_{CG} V_2 + R_G \beta_{BG}\left(\frac{V_B - \mu_{CB} V_2}{R_B + Z_B}\right)}{(R_G y_G + 1)(R_B + Z_B) + \mu_{GB} R_G \beta_{BG}}\right] +$$

$$g_{GC}(R_B + Z_B)\left[\frac{V_G + R_G g_{CG} V_2 + R_G \beta_{BG}\left(\frac{V_B - \mu_{CB} V_2}{R_B + Z_B}\right)}{(R_G y_G + 1)(R_B + Z_B) + \mu_{GB} R_G \beta_{BG}}\right]$$

$$
Y_{OUT} = \frac{I_2}{V_2} = y_C + \left[\frac{\beta_{BC}\left(\frac{V_B}{V_2} - \mu_{CB}\right)}{R_B + Z_B}\right] + \left[g_{GC}\left(R_B + Z_B\right) - \beta_{BC}\mu_{GB}\right] \cdot
$$

$$
\left[\frac{\frac{V_G}{V_2} + R_G g_{CG} + R_G \beta_{BG}\left(\frac{\frac{V_B}{V_3} - \mu_{CB}}{R_B + Z_B}\right)}{\left(R_G y_G + 1\right)\left(R_B + Z_B\right) + \mu_{GB} R_G \beta_{BG}}\right]
$$

APPENDIX IV

DERIVATION OF THE VOLTAGE GAIN FROM GRID TO COLLECTOR

IN THE COMMON EMITTER CONFIGURATION

$A_v = \dfrac{V_2}{V_3}$ Refer to figure 18 $\qquad V_2 = I_2\, g_2 \qquad g_2 = R_L/(R_L y_c + 1)$

$I_2 = (\beta_{BC} I_1 + g_{GC} V_3 \qquad\qquad I = \dfrac{V_B - \mu_{CB} V_2 - \mu_{GB} V_3}{R_B + Z_B}$

$$V_2 = \frac{(\beta_{BC} I_1 + g_{GC} V_3)\, R_L}{R_L y_c + 1}$$

$$V_2 = -\left[\frac{\beta_{BC}(V_B - \mu_{CB} V_2 - \mu_{GB} V_3)}{R_B + Z_B} + g_{GC} V_3\right]\frac{R_L}{R_L y_c + 1}$$

$$= -\frac{\beta_{BC} V_B R_L}{(R_L y_c + 1)(R_B + Z_B)} + \frac{\beta_{BC}\mu_{CB} R_L V_2}{(R_L y_c + 1)(R_B + Z_B)} + \frac{\beta_{BC}\mu_{GB} R_L V_3}{(R_L y_c + 1)(R_B + Z_B} - \frac{g_{GC} R_L V_3 (R_B + Z_B)}{(R_L y_c + 1)(R_B + Z_B)}$$

$$\left[\frac{(R_B + Z_B)(R_L y_c + 1) - \beta_{BC}\mu_{CB} R_L}{(R_B + Z_B)(R_L y_c + 1)}\right] V_2 = \left[\frac{\beta_{BC}\mu_{GB} R_L - g_{GC} R_L (R_B + Z_B)}{(R_B + Z_B)(R_L y_c + 1)}\right] V_3 - \frac{\beta_{BC} R_L V_B}{(R_B + Z_B)(R_L y_c + 1)}$$

$$A_v = \left[\frac{\beta_{BC}\mu_{GB} R_L - g_{GC} R_L (R_B + Z_B)}{(R_B + Z_B)(R_L y_c + 1) - \beta_{BC}\mu_{CB} R_L}\right] - \left[\frac{\beta_{BC} R_L}{(R_B + Z_B)(R_L y_c + 1) - \beta_{BC}\mu_{CB} R_L}\right]\frac{V_B}{V_3}$$

APPENDIX V

DERIVATION OF THE CURRENT GAIN FROM BASE TO COLLECTOR IN THE COMMON EMITTER CONFIGURATION

$A_i = \dfrac{I_2}{I_1}$ Refer to figure 18 $\qquad I_2 = \beta_{BC} I_1 + g_{GC} V_3$

$$I_3 = \frac{V_G}{R_G} + g_{CG} V_2 + \beta_{BG} I_1 \qquad V_2 = I_2 z_2 \qquad z_2 = R_L/(R_L y_C + 1)$$

$$V_3 = \frac{V_G + g_{CG} R_G V_2 + \beta_{BG} R_G I_1}{R_G y_G + 1}$$

$$I_2 = \beta_{BC} I_1 + g_{GC}\left[\frac{V_G + \left(\frac{g_{CG} R_G R_L}{R_L y_C + 1}\right) I_2 + \beta_{BG} R_G I_1}{R_G y_G + 1}\right]$$

$$\left[\frac{(R_G y_G + 1)(R_L y_C + 1) - g_{GC} g_{CG} R_G R_L}{(R_G y_G + 1)(R_L y_C + 1)}\right] I_2 = \left[\frac{\beta_{BC}(R_G y_G + 1) + g_{GC}\beta_{BG} R_G}{R_G y_G + 1}\right] I_1 + \left[\frac{g_{GC} V_G}{R_G y_G + 1}\right]$$

$$A_i = \left[\frac{\{\beta_{BC}(R_G y_G + 1) + g_{GC}\beta_{BG} R_G\}(R_L y_C + 1)}{(R_G y_G + 1)(R_L y_C + 1) - g_{CG} R_G R_L g_{GC}}\right] + \left[\frac{g_{GC}(R_G y_G + 1)}{(R_G y_G + 1)(R_L y_C + 1) - g_{GC} g_{CG} R_G R_L}\right]\frac{V_G}{I_1}$$